

Db 271 GLDGDVVPVKNGSSEKRVLKDLILSPTIAMRHILIAIGIHFFQQSSGIDAVVLYSP 330
QY 306 RLFKSAGITDDNKLGLGVTCAVGVTKEFFILVATFLLDRAGRRPLLLISTGGMIVSLICLG 365
Db 331 LVFKSAGITGDSRLRGTTAVGATNTVFILVATFLLDRIRRLPLVLTSTGGMVLSVLGLA 390
QY 366 SGLTVAGHHPDKVAVAWALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFAVG 425
Db 391 TGLTVISRHPDEKITWAIIVLCIFCINMAYVAFFSIGLGPITWVYSSEIFPLHVRALGCSLG 450
QY 426 VASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGTLEEMGKL 485
Db 451 VAVNRLTSGVISMTFISLSKAMTIGGAPFLFAGIASFAWVFFFAVLPETRGTLEDMSL 510
QY 486 FGMPDT---GVAEEAEADAAKEKVVELPSS 512
Db 511 FGNTATHKQGAEEADDDAG--EKKVEMAAT 538

RESULT 8
ABU08338 standard; protein; 539 AA.
XX AC ABU08338;
XX DT 29-MAY-2003 (first entry)
XX DE Wheat sugar transport protein #4.
XX KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.
XX OS Triticum aestivum.
XX PN US2002178468-A1.
XX PD 28-NOV-2002.
XX PF 17-JAN-2002; 2002US-00051902.
XX PR 24-APR-1998; 98US-0083044P.
XX PR 14-APR-1999; 99US-00291922.
XX PA (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2003-340957/32.
XX DR N-PSDB; ABX93210.
XX PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX PS Claim 10; Fig 2; 56pp; English.
XX CC The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08334-
CC ABU08339 represent Beta vulgaris-like sugar transport proteins

SQ Sequence 539 AA;
Query Match 69.0%; Score 1764.5; DB 6; Length 539;
Best Local Similarity 68.8%; Pred. No. 6.3e-169;
Matches 351; Conservative 63; Mismatches 91; Indels 5; Gaps 2;
QY 6 LAKAVEPRKKGNVYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQLEILIG 65
Db 31 LPAAVEPRKKGNVRFACAILASMTSILLGYDYGVMGSAASYIQKDLKINDTQLEVLGM 90
QY 66 ILSLYSLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGV 125
Db 91 ILNVYSLIGSFAAGRTSDWIGRRFTIVFAAVIFFAGALIMGFSVNYAMLMEGRFVAGIV 150
QY 126 GYGGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLYLSNFAFARLPLHLGWRVMLAI 185
Db 151 GYALMIAPVNTGEVSPASARGVLTSFPEVFINFGILLGYVSNFAFARLSRLGWRIMLGI 210
QY 186 GAVPSGLLALLVFCMPESPRLWLKGRADARAVLEKTSATPEEAERLADIKAAAGIPK 245
Db 211 GAVPSVLLAFMVLGMPESPRLWLKGRADAKVLAKTSDTPEEAERIADIKTAAGIPL 270
QY 246 GLDGDVTVPGKEGGGELQVWKLLILSPTPAVRRILLSAVGLHFFQQASGSDSVVQYSA 305
Db 271 GLDGDVVPVKNGSSEKRVLKDLILSPTIAMRHILIAIGIHFFQQSSGIDAVVLYSP 330
QY 306 RLFKSAGITDDNKLGLGVTCAVGVTKEFFILVATFLLDRAGRRPLLLISTGGMIVSLICLG 365
Db 331 LVFKSAGITGDSRLRGTTAVGATNTVFILVATFLLDRIRRLPLVLTSTGGMVLSVLGLA 390
QY 366 SGLTVAGHHPDKVAVAWALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFAVG 425
Db 391 TGLTVISRHPDEKITWAIIVLCIFCINMAYVAFFSIGLGPITWVYSSEIFPLHVRALGCSLG 450
QY 426 VASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGTLEEMGKL 485
Db 451 VAVNRLTSGVISMTFISLSKAMTIGGAPFLFAGIASFAWVFFFAVLPETRGTLEDMSL 510
QY 486 FGMPDT---GVAEEAEADAAKEKVVELPSS 512
Db 511 FGNTATHKQGAEEADDDAG--EKKVEMAAT 538
RESULT 9
AAU97212 standard; protein; 523 AA.
XX AC AAU97212;
XX DT 27-AUG-2002 (first entry)
XX DE Soybean Beta vulgaris-like sugar transport protein.
XX KW Soybean; Beta vulgaris-like sugar transport protein;
KW carbohydrate transport; grain filling; annual field crop; plant.
XX OS Glycine max.
XX PN US6383776-B1.
XX PD 07-MAY-2002.
XX PF 14-APR-1999; 99US-00291922.
XX PR 24-APR-1998; 98US-0083044P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX WPI; 2002-453364/48.
XX DR N-PSDB; ABK51973.

PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
XX
XX
PS Example 4; Fig 2; 54pp; English.
XX
CC The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence represents a soybean Beta vulgaris-like sugar transport
CC protein
XX
SQ Sequence 523 AA;

Query Match 60.3%; Score 1542; DB 5; Length 523;
Best Local Similarity 62.0%; Pred. No. 1.8e-146;
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKGNVKYASICAILASMASVILGYDIGVMSGAAIYIK 50
Db 1 MTEGKLVAAEAHKTLDQDFPPKKRKRNYAFACAMLASMTSILLGYDIGVMSGAAIYIK 60
QY 51 KDLNITDVQLEILLIGILSYLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVN 110
Db 61 RDLKVSDEQIEILLGIINLYSLIGSCLAGRTSDWIGPRYTIIVFAGTIFVVGALLMGFSPN 120
QY 111 YGMLMAGRFVAGVGVGGMIAPVYTAEISPAASRGFLTTFPEVFINGILLGYLSNFAF 170
Db 121 YSFLMFGRFVAGIGIGYALMIAPVYTAEVSPASSRGFLTSPFEVINGILIGYISNYAF 180
QY 171 ARLPLHLGWRVMLAIGAVPSGLLALLVFCMPESPRWLVLKGLADARAVLEKTSATPEEA 230
Db 181 SKLTCLKVGRWMLGVGAIPSVLLTVGVLPAMPESPRWLVMRGLGEARKVLNKTSDSKEEA 240
QY 231 AERLADIKAAAGIPKGLDGDVTVPGKEQGGELQVWKLLILSPTPAVRRILLSAVGLHF 290
Db 241 QLRLAEIKQAAGIPESCNDDVVQVKNQSGEG--VKELFLYPTPAIRHIVIALGIHF 297
QY 291 FQASGSDSVVQYSARLFKSAGITDDNKLGVTCVAVGTCTFFILVATFLLDRAGRRPLL 350
Db 298 FQASGVDAVLYSPRIFEKAGITNDTHKLLATVAVGVKTVFILAATFTLDRVGRRPLL 357
QY 351 LISTGGMIVSLICLGSGLTVAGHHPDTKVAVAVALCIASTLSYIAFFSIGLGPITGVYTS 410
Db 358 LSSVGGMVLSLTLAISLTVI-DHSERKLMWAVGSSIAMVLAYVATFSGAGPITWVYSS 416
QY 411 EIFPLQVRALGFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFETC 470
Db 417 EIFPLRLRAQGAAGVAVNRVTSVAVSMFLSLTRAITIGGAFFLYCGIATVGWIFFYTV 476
QY 471 LPETRGRITLEMGKLF 487
Db 477 LPETRGRITLEDMEGSFG 493

RESULT 10
ABU08337
ID ABU08337 standard; protein; 523 AA.
XX
AC ABU08337;
XX
DT 29-MAY-2003 (first entry)
XX
DE Soybean sugar transport protein #3.
XX

KW Beta vulgaris-like sugar transport protein; corn; rice; wheat;
KW plant sugar transport protein; carbohydrate transport; soybean;
KW carbohydrate distribution; plant.

XX Glycine max.
OS
XX US2002178468-A1.
PN
XX 28-NOV-2002.
XX
XX 17-JAN-2002; 2002US-00051902.
XX
XX 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
XX
XX (ALLE/) ALLEN S M.
PA (HITZ/) HITZ W D.
PA (KINN/) KINNEY A J.
PA (TING/) TINGEY S V.
XX
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;
XX
XX WPI; 2003-340957/32.
DR N-PSDB; ABX93209.
XX

PT Novel plant sugar transport proteins and nucleic acid encoding the
PT protein useful for producing transgenic plants having altered levels of
PT sugar transport protein.
XX

PS Claim 10; Fig 2; 56pp; English.

XX The present invention relates to the isolation of Arabidopsis thaliana-
CC like or Beta vulgaris-like sugar transport proteins, and the
CC polynucleotide sequences encoding them. The plant sugar transport
CC proteins of the invention have been isolated from corn, rice, soybean,
CC and wheat. The polypeptides of the invention may be used for altering the
CC level of expression of a sugar transport protein in a host cell, by
CC transforming a host cell with a chimeric construct encoding all, or a
CC portion of the sugar transport protein, in sense or antisense
CC orientation. Particularly, the polypeptides may provide a means to
CC control carbohydrate transport and distribution in plants. ABU08334-
CC ABU08339 represent Beta vulgaris-like sugar transport proteins

XX Sequence 523 AA;

Query Match 60.3%; Score 1542; DB 6; Length 523;
Best Local Similarity 62.0%; Pred. No. 1.8e-146;
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKGNVKYASICAILASMASVILGYDIGVMSGAAIYIK 50
Db 1 MTEGKLVAAEAHKTLDQDFPPKKRKRNYAFACAMLASMTSILLGYDIGVMSGAAIYIK 60
QY 51 KDLNITDVQLEILLIGILSYLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVN 110
Db 61 RDLKVSDEQIEILLGIINLYSLIGSCLAGRTSDWIGPRYTIIVFAGTIFVVGALLMGFSPN 120
QY 111 YGMLMAGRFVAGVGVGGMIAPVYTAEISPAASRGFLTTFPEVFINGILLGYLSNFAF 170
Db 121 YSFLMFGRFVAGIGIGYALMIAPVYTAEVSPASSRGFLTSPFEVINGILIGYISNYAF 180
QY 171 ARLPLHLGWRVMLAIGAVPSGLLALLVFCMPESPRWLVLKGLADARAVLEKTSATPEEA 230
Db 181 SKLTCLKVGRWMLGVGAIPSVLLTVGVLPAMPESPRWLVMRGLGEARKVLNKTSDSKEEA 240
QY 231 AERLADIKAAAGIPKGLDGDVTVPGKEQGGELQVWKLLILSPTPAVRRILLSAVGLHF 290
Db 241 QLRLAEIKQAAGIPESCNDDVVQVKNQSGEG--VKELFLYPTPAIRHIVIALGIHF 297
QY 291 FQASGSDSVVQYSARLFKSAGITDDNKLGVTCVAVGTCTFFILVATFLLDRAGRRPLL 350
Db 298 FQASGVDAVLYSPRIFEKAGITNDTHKLLATVAVGVKTVFILAATFTLDRVGRRPLL 357
QY 351 LISTGGMIVSLICLGSGLTVAGHHPDTKVAVAVALCIASTLSYIAFFSIGLGPITGVYTS 410
Db 358 LSSVGGMVLSLTLAISLTVI-DHSERKLMWAVGSSIAMVLAYVATFSGAGPITWVYSS 416

PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.

PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 51.4%; Score 1316; DB 3; Length 491;		
Best Local Similarity 55.6%; Pred. No. 1.1e-123;		
Matches 264; Conservative 83; Mismatches 112; Indels 16; Gaps 4;		
QY	19 KYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQLEILIGILSLYSLFGSPAG 78	
Db	3 KFAFGCAIVASIIISIIIFGYDTGVMSGAQIFIRDDLKINDTQIEVLGILNLCALVGSILTA 62	
QY	79 ARTSDRIGRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGYGGMIAPVYTAE 138	
Db	63 GKTSDVIGRRYTIASAVIFLVGSVLNGYGNYPVLVWGRCIAGVGVGFALMIAPVYSAE 122	
QY	139 ISPAASRGELTFPPEVFINIGILLGYLSNFAFARLPHLGWRVMLAIGAVPSGLLALLVF 198	
Db	123 ISSASHRGFLTSLPELCSISGILLGYVSNYCFGKLTCLKGWRMLGIAAPPSLILAFGIT 182	
QY	199 CMPEsprwLVTKGRladARAVLEKTSATPEEAERLADIKAAAGIPKGLDGDVVTVPGKE 258	
Db	183 RMPESPRWLVMQGRLEAKKIMVLVSNTEEEAEERFRDILTAAEV-----DVTEI--KE 234	
QY	259 QGGGELQ-----VWKKLILSPTPAVRRIILLSAVGLHFFQQASGSDSVVQYSARLFKSAG 312	
Db	235 VGGGVKKKHGKXVWRELVIKPRPAVRLILIAAVGIHFFEHATGIEAVVLYSPRIFFKAG 294	
QY	313 ITDDNKLIGVTCAGVTKTFFILVATFLDRAGRRLILLISTGGMIVSLICLGSGLTVAG 372	
Db	295 VVSKDKLLLATVGVGLTKAFFIIATFLLDKVGRRKLLLTSTGGVVFALTSLAVSLTMVQ 354	
QY	373 HHPDTKVAWAVALCIASLTSLYIAFFSIGLGPITGVYTSEIFPQVRAFGAVGVASNRVT 432	
Db	355 RF--GRLAWALSLSIVSTYAFVAFSSIGLGPITWVYSSEIFPLRLRAQGAIGVAVNRIM 412	
QY	433 SAVISMTFLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRTEEMGKLF 487	
Db	413 NATVSMFSLMTKAITTGGVFFVFAGIAVAWVWVFFFMFLPETKGLPLEEMKLF 467	
RESULT 13		
AAG32071		
ID	AAG32071 standard; protein; 508 AA.	
XX		
AC	AAG32071;	
XX		
DT	17-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 38622.	
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
FD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-00301439.	
XX		
PR	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.	
PR	09-MAR-1999; 99US-01233548P.	

PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	02-JUL-1999;	99US-0142154P.
PR	06-JUL-1999;	99US-0142055P.
PR	08-JUL-1999;	99US-0142390P.
PR	09-JUL-1999;	99US-0142803P.
PR	12-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.

PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.4%; Score 1316; DB 3; Length 508;

Best Local Similarity 55.6%; Pred. No. 1.2e-123;

Matches 264; Conservative 83; Mismatches 112; Indels 16; Gaps 4;

Qy 19 KYASICAILASMASVTIGYDIGVMSGAAMYIKDNLNITDVQLEILIGILSLYSLFGSFAG 78

Db 20 KPAFGCAIVASIIISIIFGYDTGVMSGAQIFIRDDLKINDTQIEVLAGILNLCALVGSлта 79

Qy 79 ARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFBVAGVGVGGMIAPIVYTAE 138

Db 80 GKTSDVIGRRYTIALSAVIFLVGSVLMGYPNYPVLMVGRCIAGVGVGFBALMIAPVYSAE 139

Qy 139 ISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWRVMLAIGAVPSGLLALLVF 198

Db 140 ISSASHRGFLTSLPELCISLIGLVSNYCFGKTLKLGWRLMLGIAAFPSLILAFGIT 199

Qy 199 CMPESPRWLVLKGRladARAVLEKTSATPEEAERLADIKAAAGIPKGLDGVVTVPGKE 258

Db 200 RMPESPRWLVMQGRLEEAKKIMVLVSNTEEEAERFRDILTAAEV-----DVTEI--KE 251

Qy 259 QGGGELQ-----VMKKLILSPTPAVRRILLSAVGLHFFQOASGSDSVVQYSARLFKSAG 312

Db 252 VGGGVKKKHGKSVWRELVIKPRPAVRLILLAAVGIHFFEHATGIEAVVLYSPRIFKKAG 311

Qy 313 ITDDNKLLGVTCAVGVTKTFIFILVATFLLDRAGRRPLLLISTGGMIVSLICLGSGLTVAG 372

Db 312 VVSKDKLLLATVGVGLTKAFFIIITFLLDKVGRRKULLLTSTGGMVFALTSLAVSLTMVQ 371

Qy 373 HHPDTKVAVAVALCIASLTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFVGVASNRVT 432

Db 372 RF--GRLAWALSLSIVSTYAFVAFVFFSIGLGPITWVYSSEIFPLRLRAQQASIGVAVNRIM 429

Qy 433 SAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRITLSEEMKLF 487

Db 430 NATVMSFSLSMTKAITTGGVFFVFVAGIAVAAWWFFFFFMLPETKGLPLEMEKLF 484

RESULT 14

AAG32073

ID AAG32073 standard; protein; 466 AA.

XX

AC AAG32073; 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
DT 17-OCT-2000 (first entry) 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38624. 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
KW Protein identification; signal transduction pathway; metabolic pathway; 24-JUN-1999; 99US-0140695P.
KW hybridisation assay; genetic mapping; gene expression control; promoter; 28-JUN-1999; 99US-0140823P.
KW termination sequence. 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
OS Arabidopsis thaliana. 01-JUL-1999; 99US-0141842P.
XX EP1033405-A2. 01-JUL-1999; 99US-0142154P.
PN 06-SEP-2000. 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
PD 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
PF 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.

PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match		
Best Local Similarity 48.7%; Score 1245; DB 3; Length 466;		
Matches 250; Conservative 79; Mismatches 107; Indels 16; Gaps 4;		
QY	42	MSGAAMYIKKDLNITDVQLEILIGILSYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVG 101
Db	1	MSGAQIFIRDDLKINDTQIEVLAGILNLCAIVGSLTAGKTSDVICRRYTIALSAVIFLVG 60
QY	102	SLLMGFAVNYGMLMAGRFVAGVGVGGMIAPVYTAEISPAASRGFLTTFPEVFINIGIL 161
Db	61	SVLMGYGNYPVLMVGRGCIAGVGVGFALMIAPVYSAEISSASHRGFLTSLPELCISLGIL 120
QY	162	LGYLSNFAFARPLHLGWRVMLAIGAVPSGILLALLVFCMPESPRLVLKGRLLADARAVLE 221
Db	121	LGYVSNYCFGKLTLLKGWRMLMLGIAAFPSLILAFGITRMPESPRLVMQGRLEEAKIMV 180
QY	222	KTSATPEEAERLADIKAAAGIPKGLDGVVTVPGKEQGGGELQ-----VWKKLILSPT 275
Db	181	LVSNTTEBAEERFRDILTAAEV-----DVTEI--KEVGGGVKKQHGKSVWRELVIKPR 232

QY	276	PAVRILLSAVGLHFFQOASGDSVVQYSARLFKSAGITDDNKKLLGVTCAVGVTKTFFIL 335
Db	233	PAVRLILIAAVGIHFFEHATGIEAVVLYSPRIFKKAGVSVSKDLLLATVGVGLTKAFFII 292
QY	336	VATFLLDRAGRRPRLILLISTGGMIVSVLICLSGLTVAGHHPDTKVAVAVALCIASTLSYIA 395
Db	293	IATFLLDKVGRKKLLTSTGGMVFALTSLAVSLTMVQRF--GRLAWALSLSIVSTYAFVA 350
QY	396	FFSIGLGPITGVYTSEIFPPLQVRALGFAVGVASNRVTSAVISMTFLSKAITIGGSPFL 455
Db	351	FFSIGLGPITWVYSSEIFPLRLRAQGASIGVAVNRIMNATVSMFSLSMTKAITTGGVFFV 410
QY	456	YSGIAAVAWVFFFTCLPETRGRRTLEEMGKLFG 487
Db	411	FAGIAVAAWWFFFPMLPETKGLPLEEMEKLFG 442
RESULT 15		
ID	ADA48320	
XX	ADA48320 standard; protein; 333 AA.	
AC	ADA48320;	
XX		
DT	20-NOV-2003 (first entry)	
XX		
DE	Rice protein conferring disease resistance in plants.	
XX		
KW	disease resistance; pathogen tolerance; plant pathogen; plant; rice.	
XX		
OS	Oryza sativa.	
XX		
PN	WO2003000906-A2.	
XX		
PD	03-JAN-2003.	
XX		
PF	21-JUN-2002; 2002WO-IB002453.	
XX		
PR	22-JUN-2001; 2001US-0300112P.	
PR	26-SEP-2001; 2001US-0352277P.	
PR	22-MAR-2002; 2002US-0366535P.	
XX		
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX		
PI	Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;	
PI	Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;	
XX		
DR	WPI; 2003-184052/18.	
DR	N-PSDB; ADA48319.	
XX		
PT	New polynucleotide comprising a plant nucleotide sequence having an open	
PT	reading frame that encodes a polypeptide associated with disease	
PT	resistance, useful for conferring resistance or tolerance to a plant	
PT	pathogen.	
XX		
PS	Claim 10; SEQ ID NO 390; 299pp; English.	
XX		
CC	The invention relates to a novel isolated polynucleotide comprising a	
CC	plant nucleotide sequence having an open reading frame that encodes a	
CC	polypeptide associated with disease resistance or its fragment having	
CC	substantially the same activity as the full-length polypeptide. The	
CC	polynucleotide of the invention is useful for conferring resistance or	
CC	tolerance to a plant pathogen. The present sequence represents a protein	
CC	conferring disease resistance used in the invention.	
XX		
SQ	Sequence 333 AA;	
Query Match		
Best Local Similarity 30.2%; Score 772.5; DB 6; Length 333;		
Matches 174; Conservative 53; Mismatches 84; Indels 169; Gaps 6;		
QY	7	AKAVEPRKKGNVKYASICAILASMASVILGDIGVMSGAAMYIKKDLNITDVQLEILIGI 66
Db	20	AAASLPRRN---KYFFFCAVLASMTX-----RAGRRSSA-----EDLGVSDAQIEVLSGA 66

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:33:48 ; Search time 22 Seconds
(without alignments)
1203.824 Million cell updates/sec

Title: US-10-051-902A-20
Perfect score: 2559
Sequence: 1 MASDELAKAVEPRKKNVYK.....AAEAEDAAKEKVVELPSSK 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTIUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2559	100.0	513	4	US-09-291-922-20
2	1906.5	74.5	529	4	US-09-291-922-28
3	1872.5	73.2	510	4	US-09-291-922-22
4	1764.5	69.0	539	4	US-09-291-922-26
5	1542	60.3	523	4	US-09-291-922-24
6	1483.5	58.0	549	4	US-09-291-922-30
7	664.5	26.0	488	4	US-10-162-012-46
8	581.5	22.7	514	4	US-09-489-039A-11902
9	576	22.5	167	4	US-09-291-922-18
10	565	22.1	584	2	US-08-928-692-13
11	565	22.1	584	4	US-09-339-972-13
12	551.5	21.6	517	4	US-09-679-686B-18
13	542	21.2	501	4	US-09-489-039A-11731
14	533	20.8	476	4	US-09-489-039A-11933
15	528	20.6	502	4	US-09-679-686B-2
16	511	20.0	518	4	US-09-679-686B-23
17	505.5	19.8	494	2	US-09-031-392-5
18	505.5	19.8	494	3	US-09-299-549-5
19	505.5	19.8	494	4	US-09-610-417-5
20	502.5	19.6	729	4	US-09-291-922-29
21	495	19.3	519	4	US-09-679-686B-24
22	489.5	19.1	510	4	US-09-679-686B-19
23	484	18.9	493	2	US-09-031-392-10
24	484	18.9	493	3	US-09-299-549-10
25	484	18.9	493	4	US-09-610-417-10
26	482	18.8	514	4	US-09-679-686B-22
27	480.5	18.8	524	2	US-08-928-692-12

28	480.5	18.8	524	4	US-09-339-972-12	Sequence 12, Appl
29	477.5	18.7	737	4	US-09-291-922-8	Sequence 8, Appli
30	476.5	18.6	518	4	US-09-679-686B-16	Sequence 16, Appl
31	468.5	18.3	511	4	US-09-679-686B-12	Sequence 12, Appl
32	466	18.2	521	4	US-09-489-039A-9549	Sequence 9549, Ap
33	459.5	18.0	562	4	US-10-162-012-44	Sequence 44, Appl
34	453	17.7	492	2	US-08-355-844-3	Sequence 3, Appli
35	453	17.7	492	5	PCT-US95-16126-3	Sequence 3, Appli
36	453	17.7	747	4	US-09-291-922-2	Sequence 2, Appli
37	443	17.3	504	4	US-09-679-686B-21	Sequence 21, Appl
38	439.5	17.2	509	2	US-09-031-392-6	Sequence 6, Appli
39	439.5	17.2	509	3	US-09-299-549-6	Sequence 6, Appli
40	439.5	17.2	509	4	US-09-610-417-6	Sequence 6, Appli
41	421.5	16.5	534	2	US-09-031-392-4	Sequence 4, Appli
42	421.5	16.5	534	3	US-09-299-549-4	Sequence 4, Appli
43	421.5	16.5	534	4	US-09-610-417-4	Sequence 4, Appli
44	416	16.3	488	2	US-08-928-692-11	Sequence 11, Appl
45	416	16.3	488	4	US-09-339-972-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-291-922-20
; Sequence 20, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-09-291-922-20

Query Match	100.0%;	Score 2559;	DB 4;	Length 513;
Best Local Similarity	100.0%;	Pred. No. 4.5e-255;		
Matches 513;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MASDELAKAVEPRKKNVYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQL	60	
Db	1	MASDELAKAVEPRKKNVYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQL	60	
Qy	61	EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGR FV	120	
Db	61	EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGR FV	120	
Qy	121	AGVGVGYGGMIAPVYTAETSPAAASRGFLTTTPPEVFINIGILLYLSNFAPARLPLHLGWR	180	
Db	121	AGVGVGYGGMIAPVYTAETSPAAASRGFLTTTPPEVFINIGILLYLSNFAPARLPLHLGWR	180	
Qy	181	VMLAIGAVPSGLLALLVFCMPESPRWLVLKGLADARAVLEKTSATPEAAERLADIKAA	240	
Db	181	VMLAIGAVPSGLLALLVFCMPESPRWLVLKGLADARAVLEKTSATPEAAERLADIKAA	240	
Qy	241	AGIPKGLDGDVTVTPGKEQGGELQVWKLLILSPTPAVRILLSAVGLHFFQOASGSDSV	300	
Db	241	AGIPKGLDGDVTVTPGKEQGGELQVWKLLILSPTPAVRILLSAVGLHFFQOASGSDSV	300	
Qy	301	VQYSARLFKSAGITDDNKLGLVTCAGVGTKTFIFILVATFLLDRAGRRPRLIIISTGGMIVS	360	
Db	301	VQYSARLFKSAGITDDNKLGLVTCAGVGTKTFIFILVATFLLDRAGRRPRLIIISTGGMIVS	360	

QY 361 LICLSGLTVAGHHPTKVAWAVALLCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
Db 361 LICLSGLTVAGHHPTKVAWAVALLCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
QY 421 GFAVGASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRITL 480
Db 421 GFAVGASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRITL 480
QY 481 EMGKLFMPDTCMAEEAEADAAAKEKVVVELPSSK 513
Db 481 EMGKLFMPDTCMAEEAEADAAAKEKVVVELPSSK 513

RESULT 2

US-09-291-922-28
; Sequence 28, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-28

Query Match 74.5%; Score 1906.5; DB 4; Length 529;
Best Local Similarity 73.6%; Pred. No. 9.3e-188;
Matches 373; Conservative 58; Mismatches 71; Indels 5; Gaps 3;
QY 1 MASDELAK--AVEPRKKGNVYASICALASMASVILGYDIGVMSGAAMYIKDLNITDV 58
Db 19 MASAAALPEPGAVHPRNKGNFKYAFTCALCASMATIVLGYDVGMSCASLYIKRDLQITDV 78
QY 59 QLEILIGILSLYSLFGSFAGARTSDRIGRLTVFAAVIFFVGSLLMGFAVNYGMLMAGR 118
Db 79 QLEIMMGILSVYALIGSFLGARTSDWVGRRTVVFAAAIFNNGSLLMGFAVNYAMLVGR 138
QY 119 FVAGVGVGYGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLG 178
Db 139 FVTGIGVGYAIMVAPVYTPVSPASARGFLTSFTVEVFINVGILLGVSNYAFARLPLHLS 198
QY 179 WRVMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAAERLADIK 238
Db 199 WRVMLGIGAVPSALLALMVFGMPESPRLVMKGRADARAVLAKTSDTPEEAAVERLDQIK 258
QY 239 AAAGIPKGLDGDVVTPGKEQGGGELQVWKLLISPTPAVRRILLSAVGLHFFQASGSD 298
Db 259 AAAGIPRELDGDVVVMP-KTKGGQEKQVWKELIFSPTPAMRRILLAAALGIHFFQATGSD 317
QY 299 SVVQYSARLFKSAGITDDNKLKLVTCVAVGVTKTFILVATFLLDRAGRRPLLLISTGMI 358
Db 318 SVVLYSPRVFQSGAGITGDNHLLGATCAGVWKTLFILVATFQLDVRGRPLLLTSTAGML 377
QY 359 VSLICLSGLTVAGHHPTKVAWAVALLCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVR 418
Db 378 ACLIGLGTGLTVVGRHPDAKVPWAIGLCIVSLAYVFFSIGLPLTSVYTSEVFFPLRVR 437
QY 419 ALGFVAVGASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGR 478
Db 438 ALGFALGTSCNRVTSAAVSMFLSLSKAITIGGSFFLYAGIAAIGWIFFFTFIPETRGLP 497

QY 479 LEEMGKLFMPDTCMAEEAEADAAAKEK 505
Db 498 LEEIGKLFGMTDT--AVEAQDTATKDK 522
RESULT 3
US-09-291-922-22
; Sequence 22, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-09-291-922-22

Query Match 73.2%; Score 1872.5; DB 4; Length 510;
Best Local Similarity 72.7%; Pred. No. 2.8e-184;
Matches 372; Conservative 60; Mismatches 77; Indels 3; Gaps 2;
QY 1 MASDELAKAVEPRKKGNVYASICALASMASVILGYDIGVMSGAAMYIKDLNITDVQL 60
Db 1 MASAAALPEAVAPKPKGNVRFAFACAILASMTSILLGYDIGVMSGASLYIKDFNISDGKV 60
QY 61 EILIGILSLYSLFGSFAGARTSDRIGRLTVFAAVIFFVGSLLMGFAVNYGMLMAGR 120
Db 61 EVLMGILNLVSLIGSFAAGRTSDWIGRRYTVFAAVIFFAGXFLMGFAVNYAMLFGRFV 120
QY 121 AGVGVGYGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWR 180
Db 121 AGIGVGYALMIAPVYTAEVSPASARGFLTSFPEVFINFGILLGVSNYAFSRLPLNLGWR 180
QY 181 VMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAAERLADIKAA 240
Db 181 IMLGIGAAPSLLALMVFGMPESPRLVMKGRADAKVLEKTSATPEEAAERLADIKAA 240
QY 241 AGIPKGLDGDVVTPGKEQGGGELQVWKLLISPTPAVRRILLSAVGLHFFQASGSDSV 300
Db 241 AGIPEELDGDVVTP-KRGSNGEKRVWKELISPTPAVRRILLSGIGHFFQHALGHSV 299
QY 301 VOYSARLFKSAGITDDNKLKLVTCVAVGVTKTFILVATFLLDRAGRRPLLLISTGMI 360
Db 300 VFYSPLVFKSPGLTNDKHLGTTWPFVTKRLLFILLATFFIDVGVRPLLLGSTGGIILS 359
QY 361 LICLSGLTVAGHHPTKVAWAVALLCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
Db 360 LIGLGAGLTVVQHPDAKIPWAIGLSIASFLAYVAFSIGLGPITWYSSEIFPLQVRAL 419
QY 421 GFAVGASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRITL 480
Db 420 GCSLGVAAANRVTSGLVSMFLSLSKAITIGGSFFLYSGIAALAWVFFVYTLPETRGRITL 479
QY 481 EMGKLFMPDTCMAEEAEADAAAKEKVVVELPSS 512
Db 480 EMSKLF--DTAAASESDEPAKEKKKVVEMAAT 509

RESULT 4

US-09-291-922-26
; Sequence 26, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-291-922-26

Query Match 69.0%; Score 1764.5; DB 4; Length 539;
Best Local Similarity 68.8%; Pred. No. 4.2e-173;
Matches 351; Conservative 63; Mismatches 91; Indels 5; Gaps 2;

QY 6 LAKAVEPRKKGNVYASICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQLEILIG 65
Db 31 LPAAVEPKKGNVRFACAILASMTSILLGYDIGVMSGASLYIQKDLKINDTQLEVLNG 90

QY 66 ILSLSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFVAVNYMLMAGRFVAGVGV 125
Db 91 ILNVYSLIGSFAAGRTSDWIGRRFTIVFAAVIFFAGALIMGFSVNYAMLMPGRFVAGIV 150

QY 126 GYGGMIAPVYTAEISPAASRGFLTTPPEVFINIGILLGYLSNFAPARLPLHLGWRVMLAI 185
Db 151 GYALMIAPVNTGEVSPASARGVLTSPPEVFINIFGILLGYVSNFAFARLSRLIGWRIMLGI 210

QY 186 GAVPSGLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEAAERLADIKAAAGIPK 245
Db 211 GAVPSVLLAFMVLGMPESPRLVMKGRLLADAKVVLAKTSDTPEEAAERLADIKTAAGIPL 270

QY 246 GLDGDVTVTPGKEQGGGLQVWKLLILSPTPAVRRIILLSAVGLHFFQQASGSDSVVQYSA 305
Db 271 GLDGDVVPVPKNGSSEKRVLKDILSPTIAMRHILLIAGIGIHFFQQSSGIDAVVLYSP 330

QY 306 RLFSAGITDDNKLGLVTCAGVTKTFFFIILVATFLLDRAGRRPRLLLISIGGMIVSLICLG 365
Db 331 LVFKSAGITGDSRLRGITVAVGATNTVFIILVATFLLDRIRRRPLVLTSTCGMLVSLVGLA 390

QY 366 SGLTVAGHHPDTKVAVAVALCIASLTSLYIAFFSICLGPITGVYTSEIFPLQVRALGFVAG 425
Db 391 TGLTVISRHPDEKITWAILVICIFCIMAIVAVAFFSICLGPITWVYSSEIFPLHVRALGCSLG 450

QY 426 VASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGRTEEMGKL 485
Db 451 VAVNRLTSGVISMTFLSLSKAMTIGGAFFLFAGIASFAWVFFFAVYLPETRGRTEEDMSSL 510

QY 486 FGMPDT---GMAEEAEADAAAKEKVVWELPSS 512
Db 511 FGNTATHKQGAEEADDDAG--EKKVEMAAT 538

RESULT 5
US-09-291-922-24
; Sequence 24, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins

FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
US-09-291-922-24

Query Match 60.3%; Score 1542; DB 4; Length 523;
Best Local Similarity 62.0%; Pred. No. 3.6e-150;
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKGNVKYASICAILASMASVILGYDIGVMSGAAMYIK 50
Db 1 MTEGKLVAAEAHKTLODFDPKPKRKNKYAFACAMLASMTSILLGYDIGVMSGAAMYIK 60

QY 51 KDLNITDVQLEILIGILSLYSLFSGFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVN 110
Db 61 RDLKVSDEQIEILLGIINLYSLGSLAGRTSDWIGPRYTTIVFAGTIFPVGALLMGFSPN 120

QY 111 YGMLMAGRFVAGVGVGGGMIAPVYTAEISPAASRGFLTTPPEVFINIGILLGYLSNPAF 170
Db 121 YSFLMFGRFVAGIGIGYALMIAPVYTAEVSPASSRGFLTSPPEVFINGGILIGYISNPAF 180

QY 171 ARPLHLGWRVMLAIGAVPSGLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEA 230
Db 181 SKTLKVGWRMMLGVGAIPSVLLTVGLAMPESPRLVMRGLGEARKVLNKTSDSKEEA 240

QY 231 AERLADIKAAAGIPKGLDGDVTVTPGKEQGGGLQVWKLLILSPTPAVRRIILLSAVGLHF 290
Db 241 QLRLAEIKQAAAGIPESCNDVQVQVKNQSGEG---VMKELFLYPTPAIRHIVIAALGIHF 297

QY 291 FQASGSDSVVQYSARLFKSAGITDDNKLGLVTCAGVTKTFFFIILVATFLLDRAGRRPLL 350
Db 298 FQASGVDVAVWLYSPRIFERKAGITNDTHKLLATVAVGFKTVFILAATFTLDRVGRPPL 357

QY 351 LISTGGMIVSLICLGSGLTVAGHHPDTKVAVAVALCIASLTSLYIAFFSICLGPITGVYTS 410
Db 358 LSSVGGMVLSLLTLAISLTVI-DHSERKLMWAVGSSIAMVLAIVATFISIGAGPITWYSS 416

QY 411 EIFPLQVRALGFVAVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTC 470
Db 417 EIFPLRLAQGAAGVAVNRVTSVAVSMTFLSLTRAITTIGGAFFLYCGIATVGNITFFYTV 476

QY 471 LPETRGRTEEMGKLF 487
Db 477 LPETRGRTEEDMEGSFG 493

RESULT 6
US-09-291-922-30
; Sequence 30, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 549


```
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-291-922-30

Query Match      58.0%; Score 1483.5; DB 4; Length 549;
Best Local Similarity 61.6%; Pred. No. 4.2e-144;
Matches 299; Conservative 63; Mismatches 112; Indels 11; Gaps 5;

QY      4 DELAKAVEPRKXGNVYKASICAILASMASVILGYDYGVMGAAVYKDKDLNITDVQLEIL 63
Db      25 DPLKK--PPKRN--KFAFACATLASMTSVLLGYDYGVMGAAIYIKEDWHISDTQIGVL 79

QY      64 IGILSYLSFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGV 123
Db      80 VGILNIYCLFGSFAAGRTSDWIGRRYTVLAGAIFVFGALLMGFATNYAFLMVGRFVTGI 139

QY      124 GVGYGGMIAPVYTAETISPAASRGFLTTFEVEFINIGILLGYLSNFAFARLPLHLGWRVML 183
Db      140 GVGYALMIAPVYTAETISPAASRGFLTTFEVEFINAGILLGYISNLAFFSLPTHLNWRFML 199

QY      184 AIGAVPSGLLALLVFCMPESPRLWLKGRLDARAVLEKTSATPEEAERLADIKAAAGI 243
Db      200 GIGAIPSIFLAIGVLAMPESPRLWLVMQGRLDGAKKVLNRISSPEEAQLRLSEIKQTAGI 259

QY      244 PKGLDGDVTVPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDSVVQY 303
Db      260 PAECDEDIYKVEKTIKSGN-AVWKELFFNPTPAVRRAVIAGIGIHFFQQASGIDAVVLY 318

QY      304 SARLFKSAGITDDNKLGVTCVAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLIC 363
Db      319 SPRIFQSAGITNARKQLLATVAVGVVKTFLFILVATFQLDKYGRRRPLLLTSVGGMIITLT 378

QY      364 LGSGLTVA--GHHPTKVAVAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALG 421
Db      379 LAMSLTVIDHSHH---KITWAIALCITMVCVAVVASFSIGLGPITWVYSSEVFPRLRAQG 435

QY      422 FAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGTLEE 481
Db      436 TSMGVAVNRVSVGVISIFFLPLSHKITTGGAFFLFGGIAIAIAWFFFTFLPETRGTLEN 495

QY      482 MGKLF 486
Db      496 MHELF 500

RESULT 7
US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-162-012-46

Query Match      26.0%; Score 664.5; DB 4; Length 488;
Best Local Similarity 34.9%; Pred. No. 1e-59;
Matches 177; Conservative 91; Mismatches 174; Indels 65; Gaps 15;

QY      25 AILASM-ASVILGYDIGVMSG-----AAMYIKKDLNITDVQLEILIGILS 68
Db      2 ALVAALGGGFLFGYDTGVIGGFLALIDFLFRGLLTSSGALAEVLGYSTVLTGLVVSIF 61

QY      69 LYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYG-----MLMAGRFV 120
Db      62 LGRLLGSLFAGKLGDRFCRKKSLLLIALVLVFGALLSGAAPGYTTTIGLWAFYLLJIVGRVL 121

QY      121 AGVGVGYGGMIAPVYTAETISPAASRGFLTTFEVEFINIGILLGYLSNFAFARLPLH---- 176
Db      122 VGLGVGGASVLVPMYISETAPKALRGALGSLYQLAITIGILVA-----AIGLGLNKTNN 176

QY      177 -----LGWRVMLAIGAVPSGLLALLVFCMPESPRLWLKGRLDARAVLEKTSATPEEA 230
Db      177 DSALNSWGWRIPLGLQLVPALLLLIGLLFLPESPRLWVEKGLEEAREVLAKLRGV-EDV 235

QY      231 AERLADIKAAAGIPKGLDGDVTVPGKEQGGGELQVWKKLILSPT-PAVRRILLSAVGLH 289
Db      236 DQEIQEIKA-----LEATV-----SEEKAGKAS-WGELFRGRTRPKVRQRLLMGVMLQ 283

QY      290 FFOQASGSDSVVQYSARLFKSAGITDDNKLGVTCVAVGVTKTFFILVA-TFLDDRAGRRP 348
Db      284 AFQQLTGINAIFYYSPTIFKSVGVSDSVASLLVTIIVGVWNVFTFVALIFLVDRFRGRP 343

QY      349 LLLISTGGMIVSLICLGSGLTVA-----GHHPTKVAVAVALCIASTLSYIAFFSIGL 402
Db      344 LLLGAAGMAICFLILGASIGVALLLLNKPDKPSSKAAGIVA--IVFILLFIAFFALGWG 401

QY      403 PITGVYTSEIFPLQVRALGFAVGVASNRVTSAVISMTFLSLSKAI--TIGGSPFL-YSGI 459
Db      402 PIPWVILSELFTTKVRSKALATAANWLANFIIGFLFPYITGAIGLALGGYVFLVFPAGL 461

QY      460 AAVAWVFFFTCLPETRGTLEEEMGKLF 486
Db      462 LVLFILVFFVFPETKGTLEEIEELF 488
```

RESULT 8
US-09-489-039A-11902
; Sequence 11902, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11902
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11902

Query Match 22.7%; Score 581.5; DB 4; Length 514;
Best Local Similarity 29.9%; Pred. No. 3.9e-51;
Matches 147; Conservative 101; Mismatches 200; Indels 43; Gaps 9;

QY 1 MASDELAKAVEPRKKNVYAS-ICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQ 59
Db 42 MTSISNDSTLSPRTQDTRRMNWFVSIAAVAGLLFGLDIGVSGALPFTDHTFTLSSQL 101
QY 60 LEILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRF 119
Db 102 QEWVSSNMLGAIGALFNGWLSFRLGRKYSLMAGAVLFVAGSIGSAFAASVEVLLVAR 161
QY 120 VAGVGVGYGGMIAPVVYTAETISPAASRGFLTTPEVFINIGILLGYSNFAFARLPLHLG- 178
Db 162 VLGAVAGIASYTAPLYLSEMASENVRGKMISMYQLMVTGLGIVLAFSLDTAFS----YSGN 217
QY 179 WRVMLAIGAVPSGLLALLVFCMPESPRWLVLKRLADARAVLEKTSATPEEAAERLADIK 238
Db 218 WRAMLGVLALPAVILILVFLPNSPRWLAERGHIEAEVLRMLRDTSEKARDELNEIR 277
QY 239 AAAGIPKGLDGDVVTVPGKEQGGELQVWKKLILSPTPAVRILLSAVGLHFFQQAQSGSD 298
Db 278 ESLKL-----KQGG-----WALFKINRN--VRRVFLGMLLQAMQOFTGMN 316
QY 299 SVVQYSARLFKSAGITDDNKLGLVTCVAGVTKFTFFILVATFLLDRAGRPLLLISTGGM 358
Db 317 IIMYYAPRIFKMGAGTTEQQMIATLVVGLTMEATFIAVFTVDKAGRPALKIGFSVMA 376
QY 359 VSLICLG-----SGLTVAGHPDTKVAMAVALCIASLTLSYIAFFSIGLGPITGVYTS 411
Db 377 LGTLVLGYCLMQFDNGTASSG-----LSW---LSVGMTMCMCIAGYAMSAAPVVWILCSE 427
QY 412 IFPLQVRALGFVGVASNRVTSAVISMTFSLSKAITIGGSFFLYSGIAAVAVVFFFTCL 471
Db 428 IQPLKCRDFGITCSTTNWVSNMIIIGATFTLLDAIGAAGTFWLYTALNVAFIGITFWLI 487
QY 472 PETRGRITLEEM 482
Db 488 PETKNVTLEHI 498

RESULT 9
US-09-291-922-18
; Sequence 18, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins

FILE REFERENCE: BB-1163
CURRENT APPLICATION NUMBER: US/09/291,922
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 60/083,044
EARLIER FILING DATE: April 24, 1998
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Microsoft Office 97
SEQ ID NO 18
LENGTH: 167
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: UNSURE
LOCATION: (34)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (85)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (98)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (112)
FEATURE:
NAME/KEY: UNSURE
LOCATION: (151)
US-09-291-922-18

Query Match 22.5%; Score 576; DB 4; Length 167;
Best Local Similarity 73.1%; Pred. No. 2.5e-51;
Matches 114; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

QY 1 MASDELAKAVEPRKKNVYAS-ICAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQ 60
Db 10 MASAPLPAAIEPGKKNVGFACXILASMTSILLGYDIGVMSGASLYIKKDLKISDVKL 69
QY 61 EILTGILSLYSLFGSFAGARTSDRIGRRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRFV 120
Db 70 EILMGILNVYSLIGSXAAGRTSDWIGRRXTIVFAAIVFFAGAXLMGFAVNYWMLMFGRFV 129
QY 121 AGVGVGYGGMIAPVVYTAETISPAASRGFLTTTFPEVFI 156
Db 130 AGIGVGVALMIATVYTAEVSPXSARGFLTSFPEVFI 165

RESULT 10
US-08-928-692-13
; Sequence 13, Application US/08928692
; Patent No. 5958727
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,692
; FILING DATE: 12-SEPT-1997
; CLASSIFICATION: 435

US-09-679-686B-18
; Sequence 18, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB1160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-679-686B-18

Query Match 21.6%; Score 551.5; DB 4; Length 517;
Best Local Similarity 30.2%; Pred. No. 4.9e-48;
Matches 146; Conservative 90; Mismatches 196; Indels 51; Gaps 10;
QY 26 ILASMASVILGYDIGVMSGAA-----MYIKKDLNITDVQ-----LEILIG 65
Db 28 VVAATGGLIFGYDIGSVTSMNPFLLKFFPEVYDKKQMGKSASQYCKYDNQLQTFTS 87
QY 66 ILSLSYLSFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGV 125
Db 88 SLYLAALVSSFFAATVTRVGRKWSMFTGGLTFLIGAALNGAAENIAMLIVGRILLGVGV 147
QY 126 GYGGMIAPVYTAIESPAASRGFLTTFFPEVFINIGILLYLSNFAFARLPLHLGWRVMLAI 185
Db 148 GFANQSVPVYLSMAPARLRGMLNIGFQMLTIGILAAALINYDTNKIKAGYGWRISLAI 207
QY 186 GAVPSGLLALLVFCMPESPRWLVLKGRLLADARAVLEKTSATPEEAAERLADIKAAAGIPK 245
Db 208 AAVPAGIITLGSFFLDPDTPNSLIERGHPEAARRMLNRIRGSDVDISEEYADLWASE--- 264
QY 246 GLDGDVVTVPGKEQGGGELQWKKLILSPTPAVRRILLSAVGLHFFQAGSDSVVQYSA 305
Db 265 --ESKLQVHP-----WRNIL--QRKYPQLTMAIMIPFQQLTGINVIMFYAP 308
QY 306 RLKFSAGITDDNKLGVTCAGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLICLG 365
Db 309 VLFETLGFKGDAASMSAV-ITGLVNVFATLVSVFTVDRGLRRKLFLOGGTQMLLSQLVVG 367
QY 366 S-----GLTVAGHHPTDKVAVAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRA 419
Db 368 TLIAVKFGTSGVGEKPKGYAA-AVVLFC---LYVAGFAWSWGPLGWLVPSEIFPLEIRP 423
QY 420 LGFAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAFAVAVVFFTCCLPETRGRTL 479
Db 424 AGQSINVSVMMLFTFVIAQAFMLTLMCHKF-GLFYFFAGVNVIMTVFIALFLPETKNVPI 482
QY 480 EEM 482
Db 483 EEM 485

RESULT 13
US-09-489-039A-11731
; Sequence 11731, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11731
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11731

Query Match 21.2%; Score 542; DB 4; Length 501;
Best Local Similarity 29.5%; Pred. No. 4.5e-47;
Matches 148; Conservative 95; Mismatches 204; Indels 54; Gaps 12;
QY 22 SICAILASMASVILGYDIGVMSGAAVYIKKDLNITD-VQLEILIGILSLYLSFGSFAGAR 80
Db 38 TIC-LVAACGGLIFGYDWTVVIGGAKPFYEAWFSITDPAQSGWAMSSALLGCIFGALISGW 96
QY 81 TSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGMIAPVYTAIEIS 140
Db 97 CADKLGRKLLPLILSAVLSASAWGTAVASHFDMFVYRIVGGVGIGLASALSPLYIAEVS 156
QY 141 PAASRGFLTTFFPEVFINIGILLYLSNFAFARLPL-----HLGWRVMLAI 185
Db 157 PAEKRGRFVAVNQLTIVIGVLAQAQLINLMIAE-PVEPGATQQMIVDSWNGQMGWRWMFGA 215
QY 186 GAVPSGLLALLVFCMPESPRWLVLKGRLLADARAVLEKTSATPEEAAERLADIKAAAGIPK 245
Db 216 ELVPALAFVLVMPFVPESPRMLKAGKPERARAAALERTGSA--DYADRI-----LREIAH 268
QY 246 GLDGDVVTVPGKEQGGGELQWKKLILSPTPAVRRILLSAVGLHFFQAGSDSVVQYSA 305
Db 269 TLEKD---NNKVSYGALL-----APQVKPIVIGMVLAIFQQWCGINVIENYQAQ 314
QY 306 RLKFSAGITDDNKLGVTCAGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLICLG 365
Db 315 EIFASAGF-DINSTLKSIVATGVNVNLTFTIAALPLVDKIGRRKMLLGASGLTLIYV--- 370
QY 366 SGLTVAGHHPTDKVAVAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAIGFAVG 425
Db 371 ---LIAGAYAMGIMGWVLLLV--LAAIAIYALTAPVTWVLLAEIFFPNRVRLGAMSIG 424
QY 426 VASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAFAVAVVFFTCCLPETRGRILEMGKL 485
Db 425 TLALWIAACFLTYTFFPLNAGLGAAGSFLLYGVICAAGLYILRNVPTKGTITLEALEEQ 484
QY 486 FGMPDTGMAEEAEADAAAEKV 506
Db 485 LAQRHTGV-----NAAKQEQM 500

RESULT 14
US-09-489-039A-11933
; Sequence 11933, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11933
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11933

	Query Match	20.8%; Score 533; DB 4; Length 476;
	Best Local Similarity 27.9%; Pred. No. 3.5e-46;	
	Matches 136; Conservative 99; Mismatches 206; Indels 46; Gaps 9	
QY	13 RKKG--NVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQLHILIGILSLY 70	
Db	17 KKGRSNKTMTFVCFLLAALACLLFGLDIGVIAAGALPPIANEFQISAHQTQWVWSSMMFG 76	
QY	71 SLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGYGGM 130	
Db	77 AAVGAVSGWLKSFKLGRKXSLMIGAILFVAGSLFSAAPNVEILLVSRVLLGLAVGVASY 136	
QY	131 IAPVYTAEISPAASRGFLTFPEVFINIGILLGYLSNFAFARLPHLG-WRVMLAIGAVP 189	
Db	137 TAPYLSEIAPEKIRGSMISMYQLMITIGILGAYLSDTAFS---YSGAWRWMLGVIIIP 192	
QY	190 SGLLALLVFCMPESPRWLVLKGRADARAVLEKTSATPEEAAERLADIKAAAGIPKGLDG 249	
Db	193 AVLLIGVIFLPDSPRWFAAKRRFVDAERVLLRLRDTSAEAKRELDEIRSLKV----- 246	
QY	250 DVVTVPGKEQGGGELQVWKXILISPTPAVRRILLSAVGLHFFQASGSDSVVQYSARLFK 309	
Db	247 -----KQSG-----WS-LFKDNSNFRAVFLGILLQVMQQTGMNVIMYYAPKIFE 291	
QY	310 SAGITDDNKLGLVTCVAGVTKTFFILVATFLLDRAGRPRLLLISTGGMIVSLICLGSLT 369	
Db	292 LAGYANTTEQMGTIVGLTNVLATFIAIGLVDRWGRKPTLILGFIVMAAGMVGVLGTMH 351	
QY	370 VAGHHPTKVAVAWALCIASTLSYIAFFSIGLGPITGYVTSEIFPLQVRALGPAVGVASN 429	
Db	352 I-GIHSST---AQYIAVLMLLMFIVGFAMSAGPLIIVLCSEIQPLKGRDFGITCSTATN 406	
QY	430 RVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETR-----G 476	
Db	407 WIANMIVGATFLTMLNSLGSANTFWVYGGNLNVLFILLTLWLIPETKNVSLHIEHRLMQG 466	
QY	477 RTLEEMG 483	
Db	467 RPLREIG 473	

RESULT 15

```

RESUL 13
US-09-679-686B-2
; Sequence 2, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: BB160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (488)
; OTHER INFORMATION: Xaa = any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (493)
; OTHER INFORMATION: Xaa = any amino acid
US-09-679-686B-2

```

	Query Match	20.6%;	Score 528;	DB 4;	Length 502;
	Best Local Similarity	28.6%;	Pred. No. 1.3e-45;		
	Matches 140;	Conservative	90;	Mismatches 207;	Indels 52; Gaps 10;
QY	20	YASICAILASMASVILGYDIGVMGAA-----MYIKKDLNITDVQ-----	59		
Db	25	YFILACIVGSPGSLFGYDLGVSSGVTSMDDFLVKFFPDVYRRKQAHLHETDYCKYDNQV	84		
QY	60	LEILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFF	119		
Db	85	LTLFTSSLYFAGLVSTFGASYVTKRHGRRASINGGAASFFLGGAINGAAMNIAMLIIVGRI	144		
QY	120	VAGVGVGYGGMIAPVYTAESISPAASRGFLTTFFPEVFINIGILLGYLSNFAFARLPHLGW	179		
Db	145	LLGVGVGFANQAVPVYLSEMAPARLGMNLNIGFQMLTITIGILAAELINYGTNKIKAGYGV	204		
QY	180	RVMLAIGAVPSGLLALLVFCMPESPRLVLKGLADARAVLEKTSATPEEAAERLADIKA	239		
Db	205	RVSLALAAVPAAIITLGSFLPDTPNSSLERGHPEEARMLRRIRGT-DDIGEYADLVA	263		
QY	240	AAGIPKGLDGVVTVPGKEQGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDS	299		
Db	264	AS-----EEARQVRHPWRNIL--RRRYRAQLTMVAVIPFFQQLKGINV	304		
QY	300	VVOYSARLFKSAGITDDNKKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIV	359		
Db	305	IMFYAPVLFDTLGFKKAEAFILMS-SVITGLVNVFATVSVITVDVGRKKLFLOGGAQMIV	363		
QY	360	SLICLGS-----GLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIF	413		
Db	364	COLIVGTLLIAAKFGTSGTG---DIAKGYA-AVVVVFICAYVAGFAWSWGPLGLWLPSEIF	419		
QY	414	PLQVRALGFVAGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWFFFTCLPE	473		
Db	420	PLEIRPAGQSINVSVMNMFFTFCIAQAFLTMLCHKEFP-GLFYFPAGVWVIMTVFIAFFLPE	478		
QY	474	TRGRTTLEEM	482		
Db	479	TKNVPIHEEM	487		

Search completed: June 30, 2004, 18:37:51
Job time : 36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:33:08 ; Search time 50 Seconds
(without alignments)
2900.556 Million cell updates/sec

Title: US-10-051-902A-20
Perfect score: 2559
Sequence: 1 MASDELAKAVEPRKKGNVY.....AAEEADAAAKEKVLEPSSK 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09H_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2559	100.0	513	13	US-10-051-902-20
2	2559	100.0	513	13	US-10-051-909-20
3	1906.5	74.5	529	13	US-10-051-902-28
4	1906.5	74.5	529	13	US-10-051-909-28
5	1872.5	73.2	510	13	US-10-051-902-22
6	1872.5	73.2	510	13	US-10-051-909-22
7	1764.5	69.0	539	13	US-10-051-902-26
8	1764.5	69.0	539	13	US-10-051-909-26
9	1626	63.5	327	12	US-10-425-114-68399
10	1612.5	63.0	356	12	US-10-425-114-61926
11	1597	62.4	380	12	US-10-425-114-39509
12	1542	60.3	523	13	US-10-051-902-24
13	1542	60.3	523	13	US-10-051-909-24
14	1487.5	58.1	517	12	US-10-425-114-56035
15	1483.5	58.0	548	10	US-09-774-381-40

16	1483.5	58.0	549	13	US-10-051-902-30	Sequence 30, Appl
17	1483.5	58.0	549	13	US-10-051-909-30	Sequence 30, Appl
18	1468.5	57.4	541	12	US-10-425-114-66733	Sequence 66733, A
19	1441	56.3	580	12	US-10-425-114-67056	Sequence 67056, A
20	1380	53.9	546	12	US-10-423-114-63789	Sequence 63789, A
21	1299.5	50.8	424	12	US-10-424-599-199875	Sequence 199875,
22	1281.5	50.1	488	12	US-10-424-599-169603	Sequence 169603,
23	1279.5	50.0	502	12	US-10-425-114-51712	Sequence 51712, A
24	1195.5	46.7	407	12	US-10-425-114-49353	Sequence 49353, A
25	1193	46.6	531	12	US-10-424-599-182839	Sequence 182839,
26	1182	46.2	281	12	US-10-425-114-50090	Sequence 50090, A
27	1170	45.7	533	15	US-10-310-154-724	Sequence 724, App
28	1052	41.1	408	12	US-10-424-599-199163	Sequence 199163,
29	1022	39.9	417	12	US-10-425-114-49121	Sequence 49121, A
30	780.5	30.5	336	12	US-10-425-114-63429	Sequence 63429, A
31	700	27.4	249	12	US-10-424-599-190445	Sequence 190445,
32	685	26.8	231	12	US-10-425-114-50887	Sequence 50887, A
33	678	26.5	487	9	US-09-795-693-27	Sequence 27, Appl
34	678	26.5	487	13	US-10-095-139-14	Sequence 14, Appl
35	678	26.5	487	14	US-10-156-239-27	Sequence 27, Appl
36	678	26.5	487	14	US-10-199-485-27	Sequence 27, Appl
37	678	26.5	487	15	US-10-391-399-45	Sequence 45, Appl
38	668	26.1	457	15	US-10-369-493-23324	Sequence 23324, A
39	667.5	26.1	486	9	US-09-860-232A-7	Sequence 7, Appli
40	664.5	26.0	488	12	US-09-794-822-11	Sequence 11, Appl
41	664.5	26.0	488	13	US-10-094-059-4	Sequence 4, Appli
42	664.5	26.0	488	14	US-10-170-528-5	Sequence 5, Appli
43	664.5	26.0	488	14	US-10-162-012-46	Sequence 46, Appli
44	664.5	26.0	488	14	US-10-062-960B-4	Sequence 4, Appli
45	664.5	26.0	488	14	US-10-144-624-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-051-902-20
; Sequence 20, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-902-20

Query Match	100.0%;	Score	2559;	DB	13;	Length	513;
Best Local Similarity	100.0%;	Pred. No.	1.7e-230;				
Matches	513;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
Qy	1	MASDELAKAVEPRKKGNVYASICAILASMASVILGYDIGVMSGAAMYIKDLNITDVQL	60				
Db	1	MASDELAKAVEPRKKGNVYASICAILASMASVILGYDIGVMSGAAMYIKDLNITDVQL	60				
Qy	61	ELILGILSLYSLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV	120				
Db	61	ELILGILSLYSLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV	120				
Qy	121	AGVGVGYGGMIAPVYTABISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPLHLGWR	180				
Db	121	AGVGVGYGGMIAPVYTABISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPLHLGWR	180				

QY 181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLDARAVLEKTSATPEEAAERLADIKAA 240
Db 181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLDARAVLEKTSATPEEAAERLADIKAA 240
QY 241 AGIPKGLDGDVTVTPGKEQGGELQVWKKLLLSPTPAVRRILLSSAVGLHFFQOASGSDSV 300
Db 241 AGIPKGLDGDVTVTPGKEQGGELQVWKKLLLSPTPAVRRILLSSAVGLHFFQOASGSDSV 300
QY 301 VOYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360
Db 301 VOYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360
QY 361 LKLGSGLTAVAGHHPTDKVAVAVLCAIATLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
Db 361 LKLGSGLTAVAGHHPTDKVAVAVLCAIATLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
QY 421 GFAVGASNRVTSVISMVTFSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPETRGRITL 480
Db 421 GFAVGASNRVTSVISMVTFSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPETRGRITL 480
QY 481 EMGKLFMGMPDTGMAEEAEDAAAKEKVVELPSSK 513
Db 481 EMGKLFMGMPDTGMAEEAEDAAAKEKVVELPSSK 513

RESULT 2
US-10-051-909-20
; Sequence 20, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-10-051-909-20

Query Match 100.0%; Score 2559; DB 13; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.7e-230;
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASDELAKAVEPRKKNVYASICAILASMASVILGYDVGMSGAAMYIKDLNITDVQL 60
Db 1 MASDELAKAVEPRKKNVYASICAILASMASVILGYDVGMSGAAMYIKDLNITDVQL 60
QY 61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGR 120
Db 61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGR 120
QY 121 AGVGVGGMIAPIVYTAETSPASRGFLTTPEVFNIGILLGYLSNFAFARLPLHLGWR 180
Db 121 AGVGVGGMIAPIVYTAETSPASRGFLTTPEVFNIGILLGYLSNFAFARLPLHLGWR 180
QY 181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLDARAVLEKTSATPEEAAERLADIKAA 240
Db 181 VMLAIGAVPSGLLALLVFCMPESPRWLVLKGRLDARAVLEKTSATPEEAAERLADIKAA 240
QY 241 AGIPKGLDGDVTVTPGKEQGGELQVWKKLLLSPTPAVRRILLSSAVGLHFFQOASGSDSV 300
Db 241 AGIPKGLDGDVTVTPGKEQGGELQVWKKLLLSPTPAVRRILLSSAVGLHFFQOASGSDSV 300

QY 301 VOYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360
Db 301 VOYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS 360
QY 361 LKLGSGLTAVAGHHPTDKVAVAVLCAIATLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
Db 361 LKLGSGLTAVAGHHPTDKVAVAVLCAIATLSYIAFFSIGLGPITGVYTSEIFPLQVRAL 420
QY 421 GFAVGASNRVTSVISMVTFSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPETRGRITL 480
Db 421 GFAVGASNRVTSVISMVTFSLSKAITIGGSFFLYSGIAAVAVVFFTCPLPETRGRITL 480
QY 481 EMGKLFMGMPDTGMAEEAEDAAAKEKVVELPSSK 513
Db 481 EMGKLFMGMPDTGMAEEAEDAAAKEKVVELPSSK 513

RESULT 3
US-10-051-902-28
; Sequence 28, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-902-28

Query Match 74.5%; Score 1906.5; DB 13; Length 529;
Best Local Similarity 73.6%; Pred. No. 2e-169;
Matches 373; Conservative 58; Mismatches 71; Indels 5; Gaps 3;
QY 1 MASDELAK- -AVEPRKKNVYASICAILASMASVILGYDVGMSGAAMYIKDLNITDV 58
Db 19 MASALPEPGAVHPRNKGNFYAFTCALCASMATIVLGYDVGMSGASLYIKRDLQITDV 78
QY 59 QLEILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGR 118
Db 79 QLEIMMILSVYALIGSFLGARTSDWVGRRVTVVFAAAIFNNGSLLMGFAVNYAMLMVGR 138
QY 119 FVAGVGVGGMIAPIVYTAETSPASRGFLTTPEVFNIGILLGYLSNFAFARLPLHLG 178
Db 139 FVTGIGVGYAIMVAPVYTPPEVSPASRGFLTSFTTEVFINVGILLGYVSNYAFARLPLHLS 198
QY 179 WRVMLAICAVPSGLLALLVFCMPESPRWLVLKGRLDARAVLEKTSATPEEAAERLADIK 238
Db 199 WRVMLGICAVPSALLALMVFGMPESPRWLVMKGRLDARAVLAKTSDTPEEAVERLDQIK 258
QY 239 AAAGIPKGLDGDVTVTPGKEQGGELQVWKKLLLSPTPAVRRILLSSAVGLHFFQOASGSD 298
Db 259 AAAGIPRELGDVWVMP-KTKGQEKQWVKELIFSPTPAMRILLAAALGIHFFQOATGSD 317
QY 299 SVVQYSARLFKSAGITDDNKLLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGM 358
Db 318 SVVLYSPRVFQSGAGITGDNHLLGATCAMGVMKTLFILVATFQLDVRVRRPLLLTSTAGML 377
QY 359 VSLICLGSLTVAGHHPTDKVAVAVLCAIATLSYIAFFSIGLGPITGVYTSEIFPLQVR 418
Db 378 ACLIGLGTGLTVVGRHPDAKVPWAIGLCIVSILAYVSFFSIGLGPLTSVYTSEVFPPLVR 437

QY	419	ALGFAVGVASNRVTSAVISMTFLSLSKAITTIGGSFFLYSGIAAVAVVFFFTCLPETEGRT	478
		: : : : : :	
D _b	438	ALGFALGTSCNRVTSAAVSMFSLSKAITTIGGSFFLYAGIAAIGWIFFFTFIPETRGLP	497
QY	479	LEEMKLFMGPDITGMAEEAEDEAAAKEK	505
		: : : : :	
D _b	498	LEEIGKLFMGTTDT--AVEAQDSTATDKX	522

RESULT 4

```

US-10-051-909-28
; Sequence 28, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport
; FILE REFERENCE: B01163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-909-28

```

Query Match	74.5%;	Score 1906.5;	DB 13;	Length 529;
Best Local Similarity	73.6%;	Pred. No. 2e-169;		
Matches 373;	Conservative	58;	Mismatches 71;	Indels 5;
				Gaps 3;

[illegible]

RESULT 5

US-10-051-902-22
; Sequence 22, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22.
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (102)
US-10-051-902-22

Query Match	73.2%;	Score 1872.5;	DB 13;	Length 510;
Best Local Similarity	72.7%;	Pred. No. 2.9e-166;		
Matches 372;	Conservative 60;	Mismatches 77;	Indels 3;	Gaps 2;

QY	1	MASDELAKAVEPRKKGNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQL	60
Db	1	MASAAPEAVAPKKGNVRFACAILASMTSILLGYDIGVMGSAASYIKKDFNISDGKV	60
QY	61	EILIGILSYSLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMIMAGRFV	120
Db	61	EVLGMILNLYSLIGSFAAGRTSDWIGRRYTVFAAVIFFAGXFLMGFAVNYAMLMEGRFV	120
QY	121	AGVGVGCGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWR	180
Db	121	AGIGVGVALMIAPVYTAEVSPASARGFLTSPPEVFINFGILLGYVSNYAFSRPLPLNLGWR	180
QY	181	VMLAICAVPSGLLALLVFCMPESPRLVLKRLADARAVLEKTSATPEEAAERLADIKAA	240
Db	181	IMLGICAAPSVLLALMVLGMPESPRLVMKRLADAKVLEKTSDTABEAAERLADIKAA	240
QY	241	AGIPKGLDGDVTVTPGKEQGGELQVWKKLILSPTPAVRRILLSAVLHFFQOASGSDSV	300
Db	241	AGIPEELDGDVTVTP-KRGSGNEKRWKELILSPTPAMRRILLSGIGIHFFQHALGIHSV	299
QY	301	VQYSARLPKSAGITDDNKLGLVTCAGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVS	360
Db	300	VFYSPLVFKSPGLTNDKHFLGITWFFGVTKRLLFILLATFFIDGVRRLPLLLGSTGGIILS	359
QY	361	LICLSGLTVAGHHDPDKVAVALCIASTLSYIAFFSIGLGPITGVYTTSEIFPLQVRAL	420
Db	360	LICLGAGLTVVQHPDAKIPWAIGLSIASTLAYVAFFSIGLGPITWVYSSSEIFPLQVRAL	419
QY	421	GPVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRITLE	480
Db	420	GCSLGVAANRVTSVISMFTFLSLSKAITIGGSFFLYSGIAALAWVFFYTYLPETRGRITLE	479
QY	481	EMGKLFGLMPDGTGMAEEAEADAAKEKVVELPSS	512
Db	480	EMSKLFG--DTAAASESDPEAKKKKVVEMAAT	509

RESULT 6

US-10-051-909-22
; Sequence 22, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:


```

; LENGTH: 539
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-051-909-26

Query Match          69.0%; Score 1764.5; DB 13; Length 539;
Best Local Similarity 68.8%; Pred. No. 4.1e-156;
Matches 351; Conservative 63; Mismatches 91; Indels 5; Gaps 2;

QY 6 LAKAVEPRKGNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQLEILIG 65
Db 31 LPAAVEPKKGNVRFAFACAILASMTSILLGYDYGVMGASLYIQDKLKINDTQLEVLNG 90

QY 66 ILSLSLFGSPAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVG 125
Db 91 ILNVSLIGSFAAGRTSDWIGRRTIVFAAVIFFFAGALIMGFSVNYAMLMFGRFVAGIGV 150

QY 126 GYGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWRVMLAI 185
Db 151 GYALMIAPVNTGEVSPASARGVLTSFPEVFINFGILLGYVSNFAFARLSRLGLWRIMLGI 210

QY 186 GAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAAERLADIKAAAGIPK 245
Db 211 GAVPSVLLAFMVLGMPESPRLVMKGRADAKVVLAKTSDTPEEAAERIAIKTAAGIPL 270

QY 246 GLDGDVTVPGKEQGGGELQVWKLLSPTPAVRRILLSAVGLHFFQOASGDSVVQYSA 305
Db 271 GLDGDVVPVKPKGSSEKRVLKDILSPTIAMRHILLIAGIGIHFFQOSSGIDAVLYSP 330

QY 306 RLFSAGITDDNKLGVTCVAGVTKTFFILVATFLLDRAGRRPLLLLISTGGMIVSLICLG 365
Db 331 LVFKSAGITGDSRLRGTAVAGATNTVFILVATFLLDRIRRRPLVLTSTGGMIVSLVGLA 390

QY 366 SGLTVAGHHPDTKVAWAVALCIASLTSLYIAFFSIGLGPITGVYTSEIFFLQVRALGFAVG 425
Db 391 TGLTVISRHPDEKITWAVLCIFCIMAYVAFFSIGLGPITWVYSSEIFFLHVRLGCSLG 450

QY 426 VASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGRRTLEEMGKL 485
Db 451 VAVNRLTSGVISMTFISLSKAMTIGGAFFLFAGIASFVAVFFFAVYLPETRGRRTLEDMSL 510

QY 486 FGMPDT---GMAEEAEADAAAKEKVVELPSS 512
Db 511 FGNTATHKQGAEEADDDAG--EKKVEMAAT 538

RESULT 9
US-10-425-114-68399
; Sequence 68399, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 68399
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73004E08_FLI.pep
US-10-425-114-68399

Query Match          63.5%; Score 1626; DB 12; Length 327;
Best Local Similarity 99.4%; Pred. No. 1.8e-143;
Matches 326; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 1 MASDELAKAVEPRKKNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQL 60
Db 11 MASDELAKIVEPRKKNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLRITDVQL 70

QY 61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
Db 71 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFFAGSLLMGFAVNYGMLMAGRFV 130

QY 121 AGVGVGYGGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWR 180
Db 131 AGVGVGYGGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYVSNFAFARLPLHLGWR 190

QY 181 VMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAAERLADIKAA 240
Db 191 VMLAIGAVPSALLALLVFCMPESPRLVLKGRADARVVLERTSATPEEAAARLVDIKAA 250

QY 241 AGIPKGLDGDVTVTPGK-EQGGGELQVWKLLILSPTPAVRRILLSAVGLHFFQOASGSDS 299
```

```

Matches 325; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 187 AVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAAERLADIKAAAGIPKG 246
Db 1 AVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAAERLADIKAAAGIPKG 60

QY 247 LDGDVTVTPGKEQGGGELQVWKLLSPTPAVRRILLSAVGLHFFQOASGSDSVVQYSAR 306
Db 61 LDGDVTVTPNKEQGGGELQVWKLLSPTPAVRRILLSAVGLHFFQOASGSDSVVQYSAR 120

QY 307 LFKSAGITDDNKLGVTCVAGVTKTFFILVATFLLDRAGRRPLLLLISTGGMIVSLICLGS 366
Db 121 LFKSAGITDDNKLGVTCVAGVTKTFFILVATFLLDRAGRRPLLLLISTGGMIVSLICLGS 180

QY 367 GLTVAGHHPDTKVAWAVALCIASLTSLYIAFFSIGLGPITGVYTSEIFFLQVRALGFAVG 426
Db 181 GLTVAGHHPDTKVAWAVALCIASLTSLYIAFFSIGLGPITGVYTSEIFFLQVRALGFAVG 240

QY 427 ASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGRRTLEEMGKLF 486
Db 241 ASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGRRTLEEMGKLF 300

QY 487 GMPDTGMAEEAEADAAAKEKVVELPSSK 513
Db 301 GMPDTGMAEEAEADAAAKEKVVELPSSK 327

RESULT 10
US-10-425-114-61926
; Sequence 61926, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61926
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3593-006-F8_FLI.pep
US-10-425-114-61926

Query Match          63.0%; Score 1612.5; DB 12; Length 356;
Best Local Similarity 94.2%; Pred. No. 3.8e-142;
Matches 326; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 1 MASDELAKAVEPRKKNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQL 60
Db 11 MASDELAKIVEPRKKNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLRITDVQL 70

QY 61 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFV 120
Db 71 EILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFFAGSLLMGFAVNYGMLMAGRFV 130

QY 121 AGVGVGYGGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWR 180
Db 131 AGVGVGYGGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYVSNFAFARLPLHLGWR 190

QY 181 VMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEEAAERLADIKAA 240
Db 191 VMLAIGAVPSALLALLVFCMPESPRLVLKGRADARVVLERTSATPEEAAARLVDIKAA 250

QY 241 AGIPKGLDGDVTVTPGK-EQGGGELQVWKLLILSPTPAVRRILLSAVGLHFFQOASGSDS 299
```



```
Db      251 AGIPRDLGDVVAVPSKGRDGSSETQVWKQLILSPTPAVRRILLSAVGLHFFQQASGSDS 310
QY      300 VVQYSARLFKSAGITDDNKKLLGVTCVAVGVTKTFEILVATFLLDRAG 345
Db      311 VVQYSARLFKSAGITDDNKKLLGVTCVAVGVTKTFEILVATFLLDRAG 356

RESULT 11
US-10-425-114-39509
; Sequence 39509, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39509
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700239390_FLI.pep
; US-10-425-114-39509

Query Match      62.4%; Score 1597; DB 12; Length 380;
Best Local Similarity 85.4%; Pred. No. 1.2e-140;
Matches 322; Conservative 21; Mismatches 28; Indels 6; Gaps 4;

QY      139 ISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWRVMLAIGAVPSGLLALLVF 198
Db      1 ISPAASRGFLTSFPEVFINIGILLGYSNFAFARLPLRFGWRVMLGIGAAPAGLLALMVL 60
QY      199 CMPESPRWLVLKRLADARAVLEKTSATPEEAAERLADIKAAAGIPKGLDGDVVTPGKE 258
Db      61 WMPESPRWLVMKRLADARAVLEKTSASPEEAAERLADIKVAAAGIPKGLDGDVVAVP-KE 119
QY      259 QGGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQASGSDSVVQYSARLFKSAGITDDNK 318
Db      120 RSGGEMQVWRELILSPTPAIRILLSAVGLHFFQQASGSDSVVLYSPSVFKSAGITDDNK 179
QY      319 L--LGVTCAVGVTKTFEILVATFLLDRAGRRPLLLISTGGMIVSLICLSGLTVAGHHPD 376
Db      180 LLDLGVTCVAVGVAKTLFIPVATFLLDRAGRRPLLLTSTGGMIVSLVGLTGLTVVGHHPD 239
QY      377 TKVAMAVALCIASLTLSYIAFFSGLGPITGVYTSEIFPQVRLGFAVGAVASNRVTSAVI 436
Db      240 AKIPSAVALCIASLTLAYVAFFSGLGPVTGVYNSEIFPQVRLGFAVGAVACNRVTSAVI 299
QY      437 SMTFLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRTEEMGKLFQM--PDTGMA 494
Db      300 SMTFLSKGITIGGSFFLYSGIAAVGWVFFFTCLPETRGRTEEMGKLFQMPEPDTDMA 359
QY      495 EEAEADAAKEKVVLPSS 511
Db      360 -EAYNAAAKEKAVEMPA 375

RESULT 12
US-10-051-902-24
; Sequence 24, Application US/10051902
; Publication No. US20020178468A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
```

```
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/10/051,902
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
; US-10-051-902-24

Query Match      60.3%; Score 1542; DB 13; Length 523;
Best Local Similarity 62.0%; Pred. No. 2.6e-135;
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY      1 MASDELAKAVE-----PRKGNVKYASICAILASMASVILGYDIGVMSGAAIYIK 50
Db      1 MTEGKLVEAAEAHKTLDPPPKRKRKNKYAFACAMLASMTSILLGYDIGVMSGAAIYIK 60
QY      51 KDLNITDVQLEILIGILSLYSLFGSFAGARTSDRIGRRRLTVVFAAVIFFVCSLLMGFAVN 110
Db      61 RDLKVSDEQIEILLGIINLYSLIGSLAGRTSDWIGPRYTIVFAGTIFFFVGALLMGFSPN 120
QY      111 YGMLMAGRFVAGVGVGGMIAPVYTAETSPASRGFLTTFPEVFINIGILLGYSNFAF 170
Db      121 YSFLMGRFVAGIGIGYALMIAPVYTAETSPASSRGFLTSFPEVFINIGILLGYSNFAF 180
QY      171 ARLPLHLGWRVMLAIGAVPSGLLALLVFCMPESPRLVLKRLADARAVLEKTSATPEEA 230
Db      181 SKLTLKVGWRMMLGVGAIPSVLLTVGLAMPESPRWLVMRGLGEARKVLNKTSDSKEEA 240
QY      231 AERLADIKAAAGIPKGLDGDVVTPGKEQGGGELQVWKKLILSPTPAVRRILLSAVGLHF 290
Db      241 QLRLAETKQAAGIPESCNDDVVQVKNQSGEG---VWKELFLYPTPAIRHIVIAALGIHF 297
QY      291 FQASGSDSVVQYSARLFKSAGITDDNKKLLGVTCVAVGVTKTFEILVATFLLDRAGRRPLL 350
Db      298 FQASGVDVAVVLYSPRIFEKAGITNDTHKLLATVAVGVKTVFILLATFTLDRVGRRPLL 357
QY      351 LISTGMIVSLICLSGLTVAGHHPDTKVAMAVALCIASLTLSYIAFFSGLGPITGVYTS 410
Db      358 LSSVGMVLSLLTLAISLTVI-DHSERKLMWAVGSSIAMVLAIVATFSIGAGPITWVYSS 416
QY      411 EIFPLQVRLGFAVGAVASNRVTSAVISMTFLSKAITIGGSFFLYSGIAAVAVWVFFFTC 470
Db      417 EIFPLRLRAQGAAGAVAVNRVTSVAVSMTFLSLTRAITIGGAFFLYCGIATVGNWIFFYTV 476
QY      471 LPETRGRTEEMGKLF 487
Db      477 LPETRGRTELEDMEGSFG 493

RESULT 13
US-10-051-909-24
; Sequence 24, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helenjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
```



```

; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Glycine max
US-10-051-909-24

Query Match      60.3%; Score 1542; DB 13; Length 523;
Best Local Similarity 62.0%; Pred. No. 2.6e-135;
Matches 308; Conservative 70; Mismatches 105; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKGNVKYASICAILASMASVILGYDYGVMGSAAMYIK 50
Db 1 MTEGKLVAAEAHKTLDQDPKPKRKNKYAFACAMLASMTSILLGYDYGVMGSAAMYIK 60
QY 51 KDLNITDVQLEILIGILSLYSLFSGSPAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVN 110
Db 61 RDLKVSDEQIEILLGIINLYSLGSLAGRTSDWIGPRYIVFAGTIFVVGALLMGFSPN 120
QY 111 YGMLMAGRFVAGVGVGYGGMIAPVYTAEISPAASRGFLTTPPEVFINIGILLGYLSNPAF 170
Db 121 YSFLMFRFVAGIGIGYALMIAPVYTAEVSPASSRGFLTSPPEVFINGGILLIGYSNYAF 180
QY 171 ARPLHLGWRVMLAIGAVPSGLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEA 230
Db 181 SKLTLLKVGWRMMLGVGAIPSVLLTVGLAMPESPRLVMRGLGEARKVLNKTSDSKEEA 240
QY 231 AERLADIKAAAAGIPKGLDGDVVTVPGKEQGGELQVWKLLSPTPAVRRILLSAVGLHF 290
Db 241 QLRLAEIKQAAGIPESCNDVVQVKNQSGEG---VWKELFLYPTPAIRHIVIAALGIHF 297
QY 291 FQASGSDSVVQYSARLFKSAGITDDNKLKLGVTCAVGVTCTFFILVATFLDLRAGRRPLL 350
Db 298 FQASGVDAAVLYSPRIFEFKAGITNDTHKLLATVAVGFVKTVFILAAATFLDRVGRPLL 357
QY 351 LISTGGMIVSLICLGSGLTVAGHHPTDKVAMAVALCIASLTLSYIAFFSIGLGPITGVYTS 410
Db 358 LSSVGGMVLSLTLAISLTVI-DHSERKLMWAVGSSIAMVLAYVATFSIGAGPITWVYSS 416
QY 411 EIPPLQVRALGFVAVGASNRVTSAVISMTFLSLSKAITIGSGFFLYSGIAAVAVWVFFFTC 470
Db 417 EIPPLRLRAQGAAGAVAVNRRTSAVWSMTFLSLTRAITIGGAFFLYCGIATVGWIFFYTV 476
QY 471 LPETRGRTEEMGKLF 487
Db 477 LPETRGTLEDMEGSFG 493

RESULT 14
US-10-425-114-56035
; Sequence 56035, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56035
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701204606_FLI.pep
US-10-425-114-56035
```

```

Query Match      58.1%; Score 1487.5; DB 12; Length 517;
Best Local Similarity 62.3%; Pred. No. 3.2e-130;
Matches 297; Conservative 69; Mismatches 104; Indels 7; Gaps 3;

QY 11 EPRKKGNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQLEILIGILSLY 70
Db 17 KPRRN---KYAFACAILASMTSILLGYDYGVMGSAALYIQRDLKVSQVQIEILNGIINLY 73
QY 71 SLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGM 130
Db 74 SPVGSFIAGRTSDWIGRRYTIVLGAIFFFVGAILMGFSPNYAFLMFGFFAGVGIGFAFL 133
QY 131 IAPVYTAEISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPLHLGWRVMLAIGAVPS 190
Db 134 IAPVYTSEISPSRRGFLTSLPEVFLNGGILIGYISNYGFSKLALRLGWRMLMGVAIPS 193
QY 191 GLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEAAERLADIKAAAAGIPKGLDGD 250
Db 194 ILIGVAVLAMPESPRLVLKGRLLGEAKRVLYKISESEEEARLRLADIKDTAGIPQDCDD 253
QY 251 VVTVPGEQGGELQVWKLLSPTPAVRRILLSAVGLHFFQASGSDSVVQYSARLFKS 310
Db 254 VVLVSKQTHG---VWRELFLHPTPAVRHIFIASLGIHFFAQTGIDAVVLYSPRIFEX 310
QY 311 AGITDDNKLKLGVTCAVGVTCTFFILVATFLDLRAGRRPLLSTGGMIVSLICLGSGLTV 370
Db 311 AGIKSDNYRLLATVAVGFVKTVSILVATFFLDRAGRVLLCSVGLLSLLTLGLSLTV 370
QY 371 AGHHPTKVAMAVALCIASLTLSYIAFFSIGLGPITGVYTSIETFPPLQVRALGFVAVSNR 430
Db 371 V-DHSQTLNWAAGLSIAAVLSYVATFSIGSGPITWVYSSEIFPLRLRAQGAIVANR 429
QY 431 VTSAVISMTFLSLSKAITIGSGFFLYSGIAAVAVWVFFFTCLPETRGRTEEMGKLF 487
Db 430 VTSGVIAMTFLSLQKAITIGGAFFLFAGVAAVAWIFHYTLPLPETRGTLEEIEKSF 486

RESULT 15
US-09-774-381-40
; Sequence 40, Application US/09774381
; Publication No. US20030082677A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Pan, Yang
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NOVEL EDIRF, MTR-1, LSP-1, TAP-1, AND PA-I MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MNI-107CP2
; CURRENT APPLICATION NUMBER: US/09/774,381
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 08/941,354
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/010,674
; PRIOR FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: 60/061,149
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/014,347
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 60/061,159
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/474,151
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 09/004,206
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/061,143
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 09/483,414
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 09/213,571
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,890
```

```
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-774-381-40

Query Match      58.0%; Score 1483.5; DB 10; Length 548;
Best Local Similarity 61.6%; Pred. No. 8.3e-130;
Matches 299; Conservative 63; Mismatches 112; Indels 11; Gaps 5;

QY      4 DELAKAVEPRKKGNVKYASICAILASMASVILGYDIGVMSGGAAMYIKKDLNITDVQLEIL 63
Db      24 DPLKK--PPKRN--KFAPACATLASMTSVLLGYDIGVMSGGAI IYLKEDWHISDTQIGVL 78

QY      64 IGILSLYSLFGSFAGARTSDRIGRRLLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFBAGV 123
Db      79 VGIILNIYCLFGSFAAGRTSDWIGRRYTIIVLAGAIFFFVGALLMGFATNYAFLMVGRFVTGI 138

QY      124 GVGYGGMIAPVYTAETISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWRVML 183
Db      139 GVGYALMIAPVYTAEVSPASSRGFLTTSFPEVFINAGILLGYISNLFSSLPHTLSWRFML 198

QY      184 AIGAVPSGLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEAAERLADIKAAAGI 243
Db      199 GIGAIPIFIPLAIGVLAMPESPRLVLMQGRIGDAKKVNLNRISDSPEEAQLRLSEIKQTAGI 258

QY      244 PKGLDGDVVTVPGKEQGGGELQWKKLILSPTPAVRILLSAVGLHFFQOASGSDSWQY 303
Db      259 PAECDEDIYKVEKTKIKSGN-AWKELFFNPTPAVRRAVIAGIGIHFFQOASGIDAVVLY 317

QY      304 SARLEKSAGITDDNKLGLVTCAGVTKTFFILVATFLLDRAGRRPRLLLISIGMIVSLIC 363
Db      318 SPRIFQSAGITNARKQLLATVAVGVVKTLLFILVATFQLDKYGRRPRLLLTSVSGMIILIT 377

QY      364 LGSGLTVA--GHPDTKVAVAVALCIASTLSYIAFFSICLGPITGVYTSEIFPLQVRALG 421
Db      378 LAMSLTVIDHSHH---KITWAIALCITMVCVVAVSFSIGLGPITWVYSSEVFPRLRAQG 434

QY      422 FAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRTEE 481
Db      435 TSMGVAVNRVWVSGVISIFFPLPLSHKITTGGAFFLFGGIAIAIAWFFFLTFLPETRGRLEN 494

QY      482 MGKLF 486
Db      495 MHELF 499
```

Search completed: June 30, 2004, 18:37:03
Job time : 51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:30:53 ; Search time 21 Seconds
(without alignments)
2349.822 Million cell updates/sec

Title: US-10-051-902A-20
Perfect score: 2559
Sequence: 1 MASDELAKAVEPRKKGNVK.....AEEAEDAAAKEKVVELPSSK 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1483.5	58.0	549	2	T14606	probable sugar tra
2	1479.5	57.8	511	2	H84536	probable sugar tra
3	1463.5	57.2	511	2	A84537	probable sugar tra
4	1346	52.6	493	2	A85433	sugar transporter
5	1316	51.4	508	2	G84564	probable sugar tra
6	1176	46.0	547	2	C84593	probable sugar tra
7	668	26.1	457	2	E70070	metabolite transpo
8	629.5	24.6	580	2	D86426	hypothetical prote
9	624.5	24.4	461	2	D70073	metabolite transpo
10	606	23.7	521	2	G84864	probable membrane
11	597	23.3	582	2	F71431	hypothetical prote
12	595	23.3	560	2	T51485	sugar transporter
13	588	23.0	517	2	D96539	hypothetical prote
14	586.5	22.9	472	2	B26430	L-arabinose isomer
15	586.5	22.9	472	2	B91091	L-arabinose isomer
16	586.5	22.9	472	2	B85936	L-arabinose isomer
17	586	22.9	473	2	G69789	sugar transporter
18	586	22.9	557	2	T38125	myo-inositol trans
19	579	22.6	472	2	S47089	arabinose-proton s
20	575.5	22.5	471	2	AB0868	L-arabinose isomer
21	574.5	22.5	487	2	E96782	hypothetical prote
22	572.5	22.4	580	2	D84772	probable sugar tra
23	568.5	22.2	482	2	B69803	metabolite transpo
24	568	22.2	502	2	B70845	probable sygar tra
25	565	22.1	584	2	S69555	myo-inositol trans
26	564	22.0	606	2	T27072	hypothetical prote
27	563	22.0	612	2	B40538	myo-inositol trans
28	553	21.6	464	2	C91106	galactose-proton s
29	553	21.6	464	2	F85951	galactose-proton s

30	552.5	21.6	639	2	T23658	hypothetical prote
31	547	21.4	464	2	F65079	galactose-proton s
32	546	21.3	464	2	AC0877	galactose-proton s
33	546	21.3	575	2	T43400	myo-inositol trans
34	544	21.3	491	2	A26430	xylose transport p
35	544	21.3	491	2	F91255	xylose-proton symp
36	544	21.3	491	2	B86096	xylose-proton symp
37	538	21.0	464	2	F69587	L-arabinose transp
38	526.5	20.6	613	2	T27077	hypothetical prote
39	526	20.6	468	2	S10014	glucose transport
40	523	20.4	507	2	T01844	probable sugar tra
41	522	20.4	433	2	G86812	D-xylose proton-sy
42	522	20.4	534	2	S38435	hexose transport p
43	520	20.3	516	2	T12199	monosaccharid tran
44	520	20.3	547	2	A48442	membrane transport
45	515.5	20.1	522	2	E86246	glucose transporte

ALIGNMENTS

RESULT 1

T14606
probable sugar transport protein 205 - beet
C;Species: Beta vulgaris (beet)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 01-Dec-2000
C;Accession: T14606; T14617
R;Chlou, T.J.; Bush, D.R.
Plant Physiol. 110, 511-520, 1996
A;Title: Molecular cloning, immunochemical localization to the vacuole, and expression i
A;Reference number: Z18131; MUID:96351183; PMID:8742332
A;Accession: T14606
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-549 <CHI>
A;Cross-references: EMBL:U64902; NID:g1778092; PID:g1778093
A;Accession: T14617
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-541, 'SVQV' <CH2>
A;Cross-references: EMBL:U64903; NID:g1778094; PID:g1778095
C;Genetics:
A;Note: BvcdNA-205; BvcdNA-397
C;Superfamily: glucose transport protein
C;Keywords: sugar transport

Query Match 58.0%; Score 1483.5; DB 2; Length 549;
Best Local Similarity 61.6%; Pred. No. 9.8e-101;
Matches 299; Conservative 63; Mismatches 112; Indels 11; Gaps 5;

QY	4	DELAKAVEPRKKGNVKYASICALLASMASVILGYDIGVMSGAAMYIKKDLNITDVQLEIL	63
Db	25	DPLKK---PPKRN--KFAFACATLASMTSVLLGYDIGVMSGAIYLYKEDWHISDTQIGVL	79
QY	64	IGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGV	123
Db	80	VGILNIYCLFGSFAAGRTSDWIGRRYTIIVLAGAIFVVGALLMGFATNYAFLMVGRFVTGI	139
QY	124	GVGYGGMIAPIVYTAEISPAASRGFLTTTFPEVFINIGILLGYLSNFAFARLPPLHLGWRVML	183
Db	140	GVGYALMIAPIVYTAEVSPASSRGFLTSFPEVFINAGILLGYISNLAFLSSLPHTLSWRFML	199
QY	184	AIGAVPSGLLALLVFCMPESPRWLVLKGRLLADARAVLEKTSATPEEAERLADIKAAAGI	243
Db	200	GIGAIPSIFLAIGVLAMPESPRWLVMQRLGDAXKVLNRIISDSPEEAQLRLSEIKQTAGI	259
QY	244	PKGLDGDVVTVPGEQGGELQVWKLLILSPTPAVRRILLSAVGLHFFQQASGDSVVQY	303
Db	260	PAECDEDIYKVEKTKIKSGN-AWVKELFFNPTPAVRRRAVIAGIGHFFQQASGIDAVVLY	318
QY	304	SARLFKSAGITDDNKLILGVTCAGVTKTFFILVATFLLDRAGRPLLLISTGGMIVSLIC	363
Db	319	SPRIFQSAGITNARKQLLATVAVGVWKTFLFILVATFQLDKYGRRPLLLTSVGGMIAILT	378

QY 364 LGSGLTVA--GHPDTPKVAWAVALCIASTLSYIAFFSIGLGPITGVYTTSEIFPLQVRALG 421
Db 379 LAMSLTVIDHSHH--KITWAIALCITWCAVVASFSIGLGPITWYSSSEVFPRLRAQG 435
QY 422 FAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAFAVAVFFFTCLPSTRGRTLEE 481
Db 436 TSMGVAVNRVVGVSIVISFFPLSLSHKITTGGCAFFLFGGIAIAIAWFFFTLFLPSTRGRTLEN 495
QY 482 MKLKF 486
Db 496 MHELF 500
RESULT 2
H84536
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: H84536
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <STO>
A;Cross-references: GB:AE002093; NID:g4678208; PIDN:AAD26954.1; GSPDB:GN00139
A;Gene: At2g16120
A;Map position: 2
C;Superfamily: glucose transport protein
Query Match 57.8%; Score 1479.5; DB 2; Length 511;
Best Local Similarity 58.8%; Pred. No. 1.8e-100;
Matches 292; Conservative 75; Mismatches 125; Indels 5; Gaps 2;
QY 11 EPRKGNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQLEILIGILSLY 70
Db 16 EPPRGNRSRYAFACAILASMTSIIILGYDYGVMGSAIFIKDDKLSDVQLEILMILNIY 75
QY 71 SLFGSPAGARTSDRIGRLTIVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGM 130
Db 76 SLVGSAAGRTSDWLGRRYITIVLAGAFFFCGALLMGFATNYPFIMVGRFVAGVGVYAMM 135
QY 131 IAPVYTAETSPAAARGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWRVMLAIGAVPS 190
Db 136 IAPVYTAEVAPASSRGFLTSFPEIFINIGILLGYVSNYFPAKLPEHLGWRFLMGVAVPS 195
QY 191 GLLALLVFCMPESPRLVLKRLADARAVLEKTSATPEEAAERLADIKAAAGIPKGLDGD 250
Db 196 VFLAIGVLAMPESPRLVLQGRLGDAFKVLDKTSNTKEAISRLNDIKRAVGIPDDMTDD 255
QY 251 VVTVPKQEGGELQVWKLLILSPTPAVRRIILSAVGLHFFQQAAGSDSVVQYSARLFKS 310
Db 256 VIVVPNKKSGAG--KGVWKDLLVRTPSVRHILACLGIFHFAQAQSGIDAVVLYSPITFSK 313
QY 311 AGITDDNKLGLVTCVAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLICLSGLTV 370
Db 314 AGLKSKNDQLLATVAVGVVKTFLFVVGTCVDRFGRALLLTSMGGMFFSLTALGTSLTV 373
QY 371 AGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTTSEIFPLQVRALGFAVGAVSNR 430
Db 374 INRNPGQTLKWAIGLAVTTVMTFVATFSLGAGPVTWYCSIFPVLRAQOASLGVMNLN 433
QY 431 VTSAVISMTFLSLSKAITIGGSFFLYSGIAAFAVAVFFFTCLPSTRGRTLEEMLKFGMPD 490
Db 434 LMSGIIGMTFLSLSKGLTIGGAFLLFAGVAAVAVFFFTFLPSTRGVPLEEIESLFG--- 490
QY 491 TGMAEEAEEDAAAKEKV 507
Db 491 SYSANKNNVMSKGQV 507
RESULT 4
A85433
sugar transporter like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: A85433

QY 364 LGSGLTVA--GHPDTPKVAWAVALCIASTLSYIAFFSIGLGPITGVYTTSEIFPLQVRALG 421
Db 379 LAMSLTVIDHSHH--KITWAIALCITWCAVVASFSIGLGPITWYSSSEVFPRLRAQG 435
QY 422 FAVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAFAVAVFFFTCLPSTRGRTLEE 481
Db 436 TSMGVAVNRVVGVSIVISFFPLSLSHKITTGGCAFFLFGGIAIAIAWFFFTLFLPSTRGRTLEN 495
QY 482 MKLKF 486
Db 496 MHELF 500
RESULT 2
H84536
probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: H84536
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-511 <STO>
A;Cross-references: GB:AE002093; NID:g4678208; PIDN:AAD26954.1; GSPDB:GN00139
A;Gene: At2g16120
A;Map position: 2
C;Superfamily: glucose transport protein
Query Match 57.8%; Score 1479.5; DB 2; Length 511;
Best Local Similarity 58.8%; Pred. No. 1.8e-100;
Matches 292; Conservative 75; Mismatches 125; Indels 5; Gaps 2;
QY 11 EPRKGNVKYASICAILASMASVILGYDYGVMGSAAMYIKKDLNITDVQLEILIGILSLY 70
Db 16 EPPRGNRSRYAFACAILASMTSIIILGYDYGVMGSAIFIKDDKLSDVQLEILMILNIY 75
QY 71 SLFGSPAGARTSDRIGRLTIVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGM 130
Db 76 SLVGSAAGRTSDWLGRRYITIVLAGAFFFCGALLMGFATNYPFIMVGRFVAGVGVYAMM 135
QY 131 IAPVYTAETSPAAARGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWRVMLAIGAVPS 190
Db 136 IAPVYTAEVAPASSRGFLTSFPEIFINIGILLGYVSNYFPAKLPEHLGWRFLMGVAVPS 195
QY 191 GLLALLVFCMPESPRLVLKRLADARAVLEKTSATPEEAAERLADIKAAAGIPKGLDGD 250
Db 196 VFLAIGVLAMPESPRLVLQGRLGDAFKVLDKTSNTKEAISRLNDIKRAVGIPDDMTDD 255
QY 251 VVTVPKQEGGELQVWKLLILSPTPAVRRIILSAVGLHFFQQAAGSDSVVQYSARLFKS 310
Db 256 VIVVPNKKSGAG--KGVWKDLLVRTPSVRHILACLGIFHFAQAQSGIDAVVLYSPITFSK 313
QY 311 AGITDDNKLGLVTCVAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLICLSGLTV 370
Db 314 AGLKSKNDQLLATVAVGVVKTFLFVVGTCVDRFGRALLLTSMGGMFFSLTALGTSLTV 373
QY 371 AGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTTSEIFPLQVRALGFAVGAVSNR 430
Db 374 INRNPGQTLKWAIGLAVTTVMTFVATFSLGAGPVTWYCSIFPVLRAQOASLGVMNLN 433
QY 431 VTSAVISMTFLSLSKAITIGGSFFLYSGIAAFAVAVFFFTCLPSTRGRTLEEMLKFGMPD 490
Db 434 LMSGIIGMTFLSLSKGLTIGGAFLLFAGVAAVAVFFFTFLPSTRGVPLEEIESLFG--- 490
QY 491 TGMAEEAEEDAAAKEKV 507

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: A85433
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <STO>
A;Cross-references: GB:NC_001268; NID:G7270615; PIDN:CAB80333.1; GSPDB:GN00140
C;Genetics:

A;Gene: AT4g36670
A;Map position: 4
C;Superfamily: glucose transport protein

Query Match 52.6%; Score 1346; DB 2; Length 493;
Best Local Similarity 55.5%; Pred. No. 9.6e-91;
Matches 268; Conservative 94; Mismatches 115; Indels 6; Gaps 5;

QY 6 LAKAVEPRKKGNV-KYASICAILASMASVILGYDIGVMSGAAVIKDKLNITDVQLEILI 64
Db 1 MADQISGEKPAGVNRFAALQCAIVASIVSIIFGYDTGVMSGAMVFIEEDLKTNDVQIEVLT 60
QY 65 GILSYSLFGSFAGARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVG 124
Db 61 GILNLCALVGSLLAGRTSDIIGRRYTIIVLASILFMLGSLMGWGPYPVLLSGRTAGLG 120
QY 125 VGYGMIAPVYTAEISPAASRGFLTTFPEVFINIGILGYLSNFAFARLPLHLGWRVMLA 184
Db 121 VGFALMVAVPVYSAEATASHRGILLASPLHLCSIGILGYIVNYFFSKLPMHIGWRMLG 180
QY 185 IGAVPSGLLALLVFCMPESPRWLKGRADARAVLEKTSATPEEAERLADIKAAAGI- 243
Db 181 IAAVPSLVLAFGILKMPESPRWLMQGRLEKEGKEILELVNSPEEAELRFQDIKAAAGID 240
QY 244 PKGLDGDVVTVPGKEQGGGELQVWKLLILSPTPAVRRIILSAVGLHFFFOQASGSDSVVQY 303
Db 241 PKCVD-DVVKMEGKTHGE--GVWKEILIRPTPAVRRLTALGIHFFFOHAGSIEAVLLY 297
QY 304 SARLPKSAGITDDNKLGVTCAGVTKTFFFIIVATFLLDRAGRRPRLLLISTGGMIVSLIC 363
Db 298 GPRIFKKAGITTKDKLFLVTIGVIMKTTFTATLLLDKVGRRKLLLTSGVMVIALTM 357
QY 364 LGSGLTVAGHHPDTKVAVAWALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFA 423
Db 358 LGFGLTMA-QNAGGKLAWALVLSIVAAYSFAFFSIGLGPITWVYSSEVFLKLRAGAS 416
QY 424 VGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPRTGRRTLEMG 483
Db 417 LGVAVNRVMNATVMSFSLTSAITTTGGAFFMFAGVAAVAVNFFFLPETKGSLEEIE 476
QY 484 KLF 486
Db 477 ALF 479

RESULT 5
G84564
probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84564

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: G84564
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AE002093; NID:G4218010; PIDN:AAD12218.1; GSPDB:GN00139
C;Genetics:

A;Gene: At2g18480
A;Map position: 2
C;Superfamily: glucose transport protein

Query Match 51.4%; Score 1316; DB 2; Length 508;
Best Local Similarity 55.6%; Pred. No. 1.5e-88;
Matches 264; Conservative 83; Mismatches 112; Indels 16; Gaps 4;

QY 19 KYASICAILASMASVILGYDIGVMSGAAVIKDKLNITDVQLEILIGILSYSLFGSFAG 78
Db 20 KFAFGCAIVASIIISIFGYDTGVMSGAIIFIRDDLKINDTQIEVLAILNLCALVSLTA 79
QY 79 ARTSDRIGRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGMIAPVYTAE 138
Db 80 GKTSDVIGRRYTIALSAVIFLVGSVLMGYGPNYPVLMVGRCIAGVGVGFALMIAPVYSAE 139
QY 139 ISPAASRGFLTTFPEVFINIGILGYLSNFAFARLPLHLGWRVMLAIGAVPSGLLALLVF 198
Db 140 ISSASHRGFLTSLPELCSISLIGILGYVSNYCFGKLTLLKGLWRMLGIAAPSLILAFGIT 199
QY 199 CMPEsprwLVLKGRADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGDVVTVPGKE 258
Db 200 RMPESPRWLMVQGRLEEAKKIMVLVNSNTEEEAERFRDILTAAEV-----DVTEI--KE 251
QY 259 QGGGELQ-----VWKKLILSPTPAVRRIILSAVGLHFFFOQASGSDSVVQYSARLFKSAG 312
Db 252 VGGGVKKKNHKSVMRELVIKPRPAVRLILIAAVGIHFFEHATGIEAVVLYSPRIFKKAG 311
QY 313 ITDDNKLGVTCAGVTKTFFFIIVATFLLDRAGRRPRLLLISTGGMIVSLICLGSGLTVAG 372
Db 312 VVSKDKLLLATVGVGLTKAFFIIITATFLLDKVGRRKLLLTSGGMVFALTSLAVSLTMVQ 371
QY 373 HHPDTKVAVAWALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFAVGVASNRVT 432
Db 372 RF--GRLAWALSLSIVSTYAFVAFPSIGLGPITWVYSSEIFPLRLRAQGASIGVAVNRIM 429
QY 433 SAVISMFTLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPRTGRRTLEMGKLF 487
Db 430 NATVSMFSLMTKAITTTGGVFFVFAGIAVAAWVFFFFMLPETKGLPLEEMEKLFG 484

RESULT 6
C84593

probable sugar transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 03-Jun-2002
C;Accession: C84593

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-547 <STO>
A;Cross-references: GB:AE002093; NID:G4454470; PIDN:AAD20917.1; GSPDB:GN00139
C;Genetics:

A;Gene: At2g20780
A;Map position: 2
C;Superfamily: glucose transport protein

Query Match 46.0%; Score 1176; DB 2; Length 547;
Best Local Similarity 45.7%; Pred. No. 2.8e-78;
Matches 241; Conservative 96; Mismatches 138; Indels 52; Gaps 6;

QY 8 KAVEPRKKGNVKYASICAILASMASVILGY-----DIGVMSGAA 46
Db 42 REAEARNSTRKRYVMACAFASLNNVLLGYGRFYLNPRIILLLLYFVDLQKDVGVMSGAV 101
QY 47 MYIKKDLNITDVQLEILIGILSYSLFGSFAGARTSDRIGRLTVVFAAVIFFVGSLLMG 106
A;Cross-references: GB:AE002093; NID:G4218010; PIDN:AAD12218.1; GSPDB:GN00139
C;Genetics:

QY 256 GKEQGGELQVWKLLILSPTPAVRRILLSAVGLHFFQOASGSDSVVQYSARLFKSAGITD 315
Db 252 ILEEGSSEKINMIKLCCKAKT--VRRGLIAGVGLQVQFVGINTVMYISPTIVQLAGPAS 309
QY 316 DNKLLGVTCAGVTKTFFILVATFLDRAGRRPLLLISTGGMIVSLICLGSLTVAGHH- 374
Db 310 NRTALLSLVTAGLNAFGSIISIYFIDRIGRKKLLIISLFGVILSGILTVGFYEATHA 369
QY 375 -----PDTKV-----AWAVALCI-ASTLS----- 392
Db 370 PAISSLETORFNNISCPDYKSAMNTNWDCTCLKASSPCGYCSPIGKEHPGACWISD 429
QY 393 -----YIAFFSIGLGPITGVYTSSEIFPLQVRA 419
Db 430 DSVKDLCHNENRLWYTRGCPNSFNGFWALLGLGLYIIFFSPGMGTVPWIVNSEIYPLRFRG 489
QY 420 LGFAVGASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRTL 479
Db 490 ICGGIAATANWISNLIVAQSFSLTBEAGTSWTFLIQFVVISVIALLFVWVCVPETKGMPPM 549
QY 480 EEMGKL 485
Db 550 EEIEKM 555
RESULT 9
D70073
metabolite transport protein homolog yxC - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: D70073
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D70073
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-461 <KUN>
A;Cross-references: GB:Z99124; GB:AL009126; NID:92636442; PIDN:CABL6017.1; PID:92636527
A;Experimental source: strain 168
C;Genetics:
A;Gene: yxC
C;Superfamily: glucose transport protein

QY 194 ALLVFCMPESPRWLVLKRLADARAVLEKTSATPEEAAERLADIKAAAAGIPKGLDGDWVT 253
Db 177 LIGIAFMPESPRWLVKRGSEEEARRIMNITH-DPKDIEMELAEMK----- 220
QY 254 VPGKEQGGELQVWKLLILSPTPA--VRRILLSAVGLHFFQOASGSDSVVQYSARLFKSA 311
Db 221 -----QGEAEK---KETTLGVLKAKWIRPMLLIGVLAIFQOAVGINTVIYAPTIFTKA 272
QY 312 GITDDNKLGLVTCAGVTKTFFILVATFLDRAGRRPLLLISTGGMIVSLICLGSLTVA 371
Db 273 GLGTSASALG-TMGIGILNVIMCITAMILIDRVGRKKLLIWSGVGITLSAALSGVLLTL 331
QY 372 GHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSSEIFPLQVR--ALGFAVGASN 429
Db 332 GLSAST--AWMTVVFLG--VYIVFYQATWGPVWVWVLMPELFPKARGAATGFTTLVLS- 385
QY 430 RVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAWVFFFTCLPETRGRTEEM 482
Db 386 -AANLIVSLVFPPLMLSAMGIAWVFMVFSVICLLSFFFAFYMWVPETKKSLEEI 437
RESULT 10
G84864
probable membrane transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84864
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: AB4420; MUID:20083487; PMID:10617197
A;Accession: G84864
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-521 <STO>
A;Cross-references: GB:AE002093; NID:g2289003; PIDN:AAB64332.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g43330
A;Map position: 2
C;Superfamily: glucose transport protein
Query Match 23.7%; Score 606; DB 2; Length 521;
Best Local Similarity 29.9%; Pred. No. 1.1e-36;
Matches 153; Conservative 93; Mismatches 178; Indels 88; Gaps 10;
QY 16 GNVKYASICAILASMASVILGDIGVMSGAAMYIKDLNI-----TDVQL 60
Db 27 GN-SYILGLTVTAGIGLLFGYDTGTVISGALLYIKDDFEVVKQSSFLQVYNVSSFTSSKL 85
QY 61 EILGILSLYSLFGSFAGARTSDRIGRRLLTVVFAAIVFPVGSLLMGFAVNYGMLMAGRFV 120
Db 86 ETIVSMALVGAMIGAAAGGWINDYYGRKKATLFDVVPAAAGAIWMAAPDPYVLISGRLL 145
QY 121 AGVGVGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLG-W 179
Db 146 VGLGVGASVTAPVYIAEASPSEVRGGLVSTNVLMITGGQFLSYLVNSAFTQVP---GTW 202
QY 180 RVMLAIGAVPSGLLALLVFCMPESPRWLVLKRLADARAVLEKT-----SAT 226
Db 203 RWMLGVSVPVAVIQFILMLFMPESPRWLFMKNRKAIAIQVARTYDISRLEDEIDHLSAA 262
QY 227 PEEAAERLADIKAAAGIPKGLDGDVTVVPVKEQGGELQVWKLLILSPTPAVRILLSAV 286
Db 263 EEEKQKRKT-----GYLDVFR-----SKEURLAFLAGA 292
QY 287 GLHFFQOASGSDSVVQYSARLFKSAGITDDNKLGLVTCAGVTKTFFILVATFLDRAGR 346
Db 293 GLQAFQOQTGINTVMYISPTIVQMGFHSNQLALFLSLIVAAMNAAGTVVIGIYFIDHCGR 352
QY 347 RPLLLISTGGMIVSLICLG-----SGLTVAGHHPDTKVAVAWALCIASTLSYI 394

Db 353 KKLALSSSLFGVIISLLLSVSFFKQSETSSDGL-----YGLAVLGLA---LYI 399
QY 395 AFFSIGLGPITGVYTSEIFFLQVRALGFVAVGNRVTSVAVISMTEFLSLSKAITIGGSFF 454
Db 400 VFFAPGMGPVPTVNSEIYPQYRGICGMSATVNWISNLIVAQTFLTIAEAGTGMTEFL 459
QY 455 LYSQIAAVAVWVFFFTCLPETRGTLEEMGKLF 486
Db 460 ILAGI AVLAVIFVIVFVPTQGLTFSEVEQIW 491

RESULT 11

F71431

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000

C:Accession: F71431

R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: F71431

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-582 <BEV>

A:Cross-references: GB:Z97341; NID:G2244991; PID:G2245004

C:Genetics:

A:Map position: 4COP9-4G3845

C:Superfamily: glucose transport protein

Query Match 23.3%; Score 597; DB 2; Length 582;
Best Local Similarity 27.8%; Pred. No. 5.9e-36;
Matches 160; Conservative 99; Mismatches 209; Indels 108; Gaps 11;

QY 1 MASDELAKAVEP-----RKKGNVRYASICAILASMASVILGYDIGVMSGAAVYIKDLN 54

Db 1 MVEGGIAKADKTEFTCEWRTTWKTPYIMRLALSAGIGGLLFGYDTGVISGALLFIKEDFD 60

QY 55 ITDVQLEILIGLSLY---SLFGSPAGARTSDRIGRRLTVVFAAVIFFVGSLLMGPAVNY 111

Db 61 EVDKKTWLQSTIVSMAGAVGAVGGWINDKFGRRMSILIADVFLIGAVMAFAPAP 120

QY 112 GMLMAGRFVAGVGVGYGGMIAPVYTAEISPAASRGFLTFPEVFINIGILLGVLNSFAFA 171

Db 121 WVIIVGRIFVGFVGMASMTSPLYISEASPARIRGALVSTNGLLITGGQFFSVLINLAFV 180

QY 172 RLPLHLG-WRVMLAIGAVPSGLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEA 230

Db 181 HTP---GTWRWMLGVAGVPAIVQFVLMSLPESPRWLRYKDRIAESRAILERIYPADAVE 237

QY 231 AERLADIKAAAGIPK---GLDGDVTVPGKEQGGELQVWKLLILSPTPAVRILLSAVG 287

Db 238 AEMEA-LKLSVEAEKADEAIIIGDSFSAKLKGAFG-----NPVVRGLAAGIT 283

QY 288 LHFFQASGSDSVVQYSARLFKSAGITDDNKLGLVTCVAVGVTKTFIFILVATFLLDRAGR 347

Db 284 VQVAQQFVGNVTWYYSPSIVQFAGYASNKTAMALSLLTSGNALGSIVSMFVDYGR 343

QY 348 PLLLLISTGGMIVSLICLGSGLTVAGHH-----PDTKVA----- 380

Db 344 KLMIISMFGIIACLIILATVFSQAATHAPKIDAFESRTFAPNATCSAYAPLAENAPPSR 403

QY 381 WAVALCIAS-----TLIS----- 392

Db 404 WNCMKCLRSECGFCASGVQPYAPGACVVLSDDMKATCSSRGRRTFFKDGCPKFGFLAIVF 463

QY 393 ---YIAFFSIGLGPITGVYTSEIFPLQVRALGFVAVGNRVTSVAVISMTEFLSLSKAITI 449

Db 464 LGLYIVVYAPGMGTVPWIVNSEIYPLRYRGLGGIAAVSNWVSNLVSEFSLTHALGS 523
QY 450 GGSFELYSGIAAVAVWVFFFTCLPETRGTLEEMGKL 485
Db 524 SGTFLLPAGFSTIGLFFIWLVPETKGLQFEEVEKL 559

RESULT 12

T51485

sugar transporter-like protein - Arabidopsis thaliana

N:Alternate names: protein T21H19_70

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 02-Sep-2000

C:Accession: T51485

R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000

A:Reference number: Z25394

A:Accession: T51485

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-560 <SAT>

A:Cross-references: EMBL:AL391148

A:Experimental source: cultivar Columbia; BAC clone T21H19

C:Genetics:

A:Map position: 5

A:Introns: 86/1; 124/2; 161/1; 205/2; 238/3; 276/2; 307/3; 359/1; 408/1; 432/3; 462/2; 50

A:Note: T21H19_70

C:Superfamily: glucose transport protein

Query Match 23.3%; Score 595; DB 2; Length 560;
Best Local Similarity 30.7%; Pred. No. 7.9e-36;
Matches 146; Conservative 91; Mismatches 186; Indels 52; Gaps 9;

QY 27 LASMASVILGYDIGVMSGAAVYIKDLNITDVQLEILIGLSLYSLF----- 73

Db 111 VACLGAILFGYHLGVVNGALEYLAKDLGI--AENTVLQKVMHFFTPPVNGWIVSSLL 168

QY 74 ----GSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVG 128

Db 169 AGATVGSFTGGALADKFGTRTFQDLAIPLAIGAFALCATQAQSVQTMIVGRLLAGIGIS 228

QY 129 GMIAPVYTAEISPAASRGFLTFPEVFINIGILLGVLNSFAPARLPLHLGWRVMLAIGAV 188

Db 229 SAIVPLYISEISPEIRGALGSVNQIFICIGILAALIAGLPLAANP--LWVRTMFGVAVI 286

QY 189 PSGLLALLVFCMPESPRLVLKGRLLADARAVLEKTSATPEEAERLADIKAAAGIPKGLD 248

Db 287 PSVLLAIGMAFSPESPRWLVOQGVSEAEKAI-KTLYGKERVVELVRDLASG----- 338

QY 249 GDVTVPGKEQGGELQV-WKKLILSPTPAVRILLSAVGLHFFQASGSDSVVQYSARL 307

Db 339 -----QGSSEPEAGWFDLFSS---RYWKVSVGAALFLFQQLAGINAVVYVTSV 385

QY 308 FKSAGITDDNKLGLVTCVAVGVTKTFIFILVATFLLDRAGRRLILLISTGGMIVSLICLGS 367

Db 386 FRSAGIQSD---VAASALVGASNVFGTAVASSLMDKMKRSLLLTSFGMALSMELLLSLS 442

QY 368 LTVAGHPDPTKVAWAVALCIASTLSVIAFFSIGLGPITGVYTSEIFPLQVRALGFVAVGA 427

Db 443 FTW----KALAAYSGLTAVVGTVLVLSFSLGAGFPVALLLPEIFASRIRAKAVALSIG 497

QY 428 SNRVTSAVISMTFLSLSKAITIGGSFELYSGIAAVAVWVFFFTCLPETRGTLEEM 482

Db 498 MHWISNFVIGLYFLSVVTKFGISSVYLGFAGVCVLAVLYIAGNVVETKGRSLEEI 552

RESULT 13

D96539

hypothetical protein Fl413.9 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C:Accession: D96539

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-517 <STO>
A;Cross-references: GB:AE005173; NID:g5734775; PIDN:AAD50040.1; GSPDB:GN00141

C;Genetics:
A;Gene: F14I3.9
A;Map position: 1
C;Superfamily: glucose transport protein

Query Match 23.0%; Score 588; DB 2; Length 517;
Best Local Similarity 29.8%; Pred. No. 2.3e-35;
Matches 157; Conservative 96; Mismatches 207; Indels 66; Gaps 13;

QY 16 GNVKYASICAILASMASVILGYDIGVMSGAAM---YIKKDLNITDVQ----- 59
Db 20 GVTVPVIMTCIVAAMGGLLFGYDLGISGGVTSMEEFLSKFFPEVDKQMHARRETAYCKF 79

QY 60 ----LEILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMMLM 115
Db 80 DNQLLQLFTSSLYLAALASSFVASAVTRYKRYGRKISMFGVGVAFILGSLFNATNVAMLI 139

QY 116 AGRFVAGVGVGCGMIAPVYTAEISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPL 175
Db 140 VGRLLGVGVGFANQSTPVLSEMAPAKIRGALNIGFQMAITIGILIANLINYGTSMQAK 199

QY 176 HLGWRVMLAIGAVPSGLLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLA 235
Db 200 N-GWRVSLGLAAPVAVIMVIGSFVLPDTPNSMLERGKYEQAREMLQIRGA-DNVDEEFQ 257

QY 236 DIKAAAGIPKGLDGDVVTVPGKEQGGGELQVWKILSPTPAVRRILLSAVGLHFFQOAS 295
Db 258 DLCDACEAAKKVDNP-----WKNIFF--QQAKYRPALVFCSAIPFFQOIT 299

QY 296 GSDSVVQYSARLFKSAGITDDNKLIG--VTCAVGVTKTFFILVATFLLDRAGRRPLLLIS 353
Db 300 GINVMIFYAPVLFKTLGPDADDASLISAVITGAVNVVST--LVSIYAVDRYGRRLFLEG 356

QY 354 TCGMIVSLICLGS-----GLTVAGHPDTKVAVAVA-LCIASTLSYIAFFSIGLGPITG 406
Db 357 GIQMIVSQIVVGTLLIGMKFGTGTSGTLTPATADWILAFICL-----YVAGFAWSWGPLGW 411

QY 407 VYTSEIFPLQVRALGFVAVGNRVTSAVISMTFLSLSKAITIGGSSFFLYSGIAAVAVVF 466
Db 412 LVPSEICPLEIRPAGQAINVSVNMFFFTFLIGQFFLTMLCHMKF-GLFYFFGGMVAVMTVF 470

QY 467 FFTCLPETRGRTLEEMGKLFG-----MPDTGMABEEAEDAAAKE 504
Db 471 IYFLLPETKGVPIEEMGRVWKQHPFWKRYMPDDAVIGGGEENYVKE 516

RESULT 14
B26430
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 05-Oct-1988 #sequence revision 05-Oct-1988 #text_change 01-Mar-2002
C;Accession: B26430; A28075; I40996; B65067
R;Maiden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.F.
Nature 325, 641-643, 1987
A;Title: Mammalian and bacterial sugar transport proteins are homologous.
A;Reference number: A93389; MUID:87115869; PMID:3543693
A;Accession: B26430

A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-472 <MAI>
R;Maiden, M.C.J.; Jones-Mortimer, M.C.; Henderson, P.J.F.
J. Biol. Chem. 263, 8003-8010, 1988
A;Title: The cloning, DNA sequence, and overexpression of the gene araE coding for arabi
A;Reference number: A28075; MUID:88228015; PMID:2836407

A;Accession: A28075
A;Molecule type: DNA
A;Residues: 1-472 <MA2>
A;Cross-references: GB:J03732; NID:g145320; PIDN:AAA23469.1; PID:g145321
R;Stoner, C.; Schleif, R.
J. Mol. Biol. 171, 369-381, 1983
A;Title: The araE low affinity L-arabinose transport promoter. Cloning, sequence, transc
A;Reference number: I40996; MUID:84114868; PMID:6319708
A;Accession: I40996

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-25,'Y',349,'R' <RES>
A;Cross-references: EMBL:X00272; NID:g40940; PIDN:CAA25075.1; PID:g40941
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B65067

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-472 <BLAT>
A;Cross-references: GB:AE000368; GB:U00096; NID:g2367165; PIDN:AAC75880.1; PID:g1789207;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:

A;Gene: araE
A;Map position: 61 min
C;Superfamily: glucose transport protein
C;Keywords: arabinose transport; intramolecular oxidoreductase; isomerase; membrane prot.

Query Match 22.9%; Score 586.5; DB 2; Length 472;
Best Local Similarity 29.9%; Pred. No. 2.7e-35;
Matches 147; Conservative 104; Mismatches 197; Indels 43; Gaps 9;

QY 1 MASDELAKAVEPRKKGNVKYASI-CAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQ 59
Db 1 MVTINTESALTTPRSLRDTRRNMNFVSVAAVAGLLFGLDIGVIAGALPFITDHFVLTSLR 60

QY 60 LEILIGILSLYSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMMLMACRF 119
Db 61 QEWVSSMMMLGAAIGALFNGLWSLFRGRKYSLSMAGAILFVLSIGSAFATSVEMLLIARV 120

QY 120 VAGVGVGCGMIAPVYTAEISPAASRGFLTTPPEVFINIGILLGYLSNFAFARLPHLG- 178
Db 121 VLGIAVGIASYTAPLYLSEMASENVRGKMISMVQLMVTLIGIVLAFSLDTAFS-----YSGN 176

QY 179 WRVMLAIGAVPSGLLALLVFCMPESPRLVLKGRLDARAVLEKTSATPEEAERLADIK 238
Db 177 WRAMLGVLALPAVLLIILVVFLPNSPRWLAEKGRHTEAEVLRMLRDTSEKAREELNEIR 236

QY 239 AAAGIPKGLDGDVVTVPGKEQGGGELQVWKILSPTPAVRRILLSAVGLHFFQOASGSD 298
Db 237 ESLKL-----KQGG-----WALFKINRN--VRRAVFLGMLLQAMQOFTGMN 275

QY 299 SVVQYSARLFKSAGITDDNKLIGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGM 358
Db 276 IIMYAPRIFKXMGFTTTEQQMIATLVVGLTFMFATFIAVFTVDKAGRKPAKIGFSVMA 335

QY 359 VSLICLG-----SGLTVAGHPDTKVAVAVALCIASTLSYIAFFSIGLGPITGVYTSE 411
Db 336 LGTLVLGYCLMQFDNGTASSG-----LSW---LSVGMTMMCIAGYAMSAAPVVMILCSE 386

QY 412 IPPLQVRALGFVAVGNRVTSAVISMTFLSLSKAITIGGSSFFLYSGIAAVAVVFFFTCL 471
Db 387 IQPLKCRDFGITCSTTTTNWVSNMIIGATFTLLDSIGAAGTFWLYTALNIAFVGITFWLI 446

QY 472 PETRGRTEEM 482
Db 447 PETKNVTLEHI 457

RESULT 15
B91091
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain O157:H7, substrain RIMD 05
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 31-Dec-2001
C;Accession: B91091
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; PMID:21156231; PMID:11258796
A;Accession: B91091
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA037121.1; PID:g13363170; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs3698
C;Superfamily: glucose transport protein
C;Keywords: intramolecular oxidoreductase; isomerase

Query Match 22.9%; Score 586.5; DB 2; Length 472;
Best Local Similarity 29.9%; Pred. No. 2.7e-35;
Matches 147; Conservative 104; Mismatches 197; Indels 43; Gaps 9;

Qy 1 MASDELAKAVEPRKKNVYKASI-CAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQ 59
Db 1 MVTINTESALTPRSRLDTRPMNMFVSAAVAGLLFGLDIGVIAGALPFITDHFVLTSL 60

Qy 60 LEILIGILSLYSLFGSFACARTSDRIGRRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRF 119
Db 61 QEWVSSMMLGAAIGALENGWLSFRLGRKYSLMAGAILFVLGSGISAFATSVEMLIAARV 120

Qy 120 VAGVGVGYGGMIAPVYTAETSPAAARGFLTTFPEVFINIGILLGYLSNFAFARLPLHLG- 178
Db 121 VLGIAGVGIASYTAPLYLSEMASENVRGKMISMYQLMVTILGIVLAFSLDTAFS----YSGN 176

Qy 179 WRVMLAIGAVPSGLLALLVFCMPESPRLVLKGRADARAVLEKTSATPEAAERLADIK 238
Db 177 WRAMLGVLLPAVLLIILVFLPNSPRWLAEKGRHIEABEVLRLRDTSEKAREELNEIR 236

Qy 239 AAAGIPKGLDGDVTVPGKEGGGELQVWKLLSPTPAVRRILLSAVGLHFFQQASGSD 298
Db 237 ESLKL-----KQGG-----WALFKINRN--VRAVFLGMLLQAMQQFTGMN 275

Qy 299 SVVQYSARLFKSAGITDDNKLGVTCAGVTKTFFILVATFLLDRAGRRPLLLISTGGMI 358
Db 276 IIMYAPRIFKMGAGFTTTEQQMIATLVVGLTFMFATFIAVFTVDKAGRKPALKIGFSVMA 335

Qy 359 VSLICLG-----SGLTVAGHHPTDKVAVAVALCIASLTSLYIAFFSIGLGPITGVYTSE 411
Db 336 LGTLVLGYCLMQFDNGTASSG-----LSW---LSVGMTMNCIAGYAMSAAPVVWILCSE 386

Qy 412 IFPLQVRALGFAVGVAASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWFFFTCL 471
Db 387 IQPLKCRDFGICSTTTNWNWSNMIGATFLTLLDSIGAAGTFLWLTALNIAFVGITFWLI 446

Qy 472 PETRGRTEEM 482
Db 447 PETKNVTLEHI 457

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:31:52 ; Search time 17 Seconds
(without alignments)
1571.294 Million cell updates/sec

Title: US-10-051-902A-20
Perfect score: 2559
Sequence: 1 MASDELAKAVEPRKKNVXY.....AAEEDAAAKEKVLPSSK 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	642.5	25.1	629	1 MYCT_HUMAN	Q96qe2 homo sapien
2	624.5	24.4	461	1 CSBC_BACSU	P46333 bacillus su
3	591.5	23.1	457	1 XYL1_LACBR	O52733 lactobacill
4	586.5	22.9	472	1 ARAE_ECOLI	P09830 escherichia
5	586	22.9	557	1 ITR2_SCHPO	P87110 schizosacch
6	579	22.6	472	1 ARAE_KLEOX	P45598 klebsiella
7	568.5	22.2	482	1 YFIG_BACSU	P54723 bacillus su
8	565	22.1	584	1 ITR1_YEAST	P30605 saccharomyc
9	563	22.0	612	1 ITR2_YEAST	P30606 saccharomyc
10	547	21.4	464	1 GALP_ECOLI	P37021 escherichia
11	546	21.3	575	1 ITR1_SCHPO	Q10286 schizosacch
12	544	21.3	491	1 XYLE_ECOLI	P09098 escherichia
13	538	21.0	464	1 ARAE_BACSU	P96710 bacillus su
14	526	20.6	468	1 GLCP_SYNY3	P15729 synechocyst
15	522	20.4	534	1 HUP3_CHLKE	Q39525 chlorella k
16	520	20.3	547	1 GTR1_LEIDO	Q01440 leishmania
17	515.5	20.1	477	1 GTR8_HUMAN	Q9ny64 homo sapien
18	515.5	20.1	522	1 STP1_ARATH	P23586 arabidopsis
19	509.5	19.9	477	1 GTR8_MOUSE	Q9jif3 mus musculu
20	505.5	19.8	494	1 GTR3_SHEEP	P47843 ovis aries
21	505.5	19.8	495	1 GTR3_CANFA	P47842 canis famil
22	504.5	19.7	523	1 STC_RICCO	Q41144 ricinus com
23	502.5	19.6	534	1 HUP1_CHLKE	P15686 chlorella k
24	501.5	19.6	494	1 GTR3_BOVIN	P58352 bos taurus
25	498	19.5	478	1 GTR8_RAT	Q9jjz1 rattus norv
26	489.5	19.1	493	1 GTR3_RAT	Q07647 rattus norv
27	489.5	19.1	510	1 HEX6_RICCO	Q07423 ricinus com
28	486	19.0	490	1 GTR1_CHICK	P46896 gallus gall
29	480.5	18.8	524	1 GTR2_HUMAN	P11168 homo sapien
30	476.5	18.6	523	1 GTR2_MOUSE	P14246 mus musculu
31	474	18.5	492	1 GTR1_RABIT	P13355 oryctolagus
32	473	18.5	492	1 GTR1_RAT	P11167 rattus norv
33	471	18.4	492	1 GTR1_MOUSE	P17809 mus musculu

34	470.5	18.4	493	1 GTR3_MOUSE	P32037 mus musculu
35	468	18.3	492	1 GTR1_HUMAN	P11166 homo sapien
36	464	18.1	496	1 GTR3_CHICK	P28568 gallus gall
37	463	18.1	496	1 GTR3_HUMAN	P11169 homo sapien
38	462	18.1	592	1 HXT5_YEAST	P38695 saccharomyc
39	461	18.0	492	1 GTR1_BOVIN	P27674 bos taurus
40	459.5	18.0	541	1 GT10_HUMAN	O95528 homo sapien
41	459	17.9	540	1 HUP2_CHLKE	Q39524 chlorella k
42	457.5	17.9	522	1 GTR2_RAT	P12336 rattus norv
43	451.5	17.6	522	1 STA_RICCO	Q10710 ricinus com
44	447.5	17.5	509	1 GTR4_MOUSE	P14142 mus musculu
45	447.5	17.5	533	1 QUTD_EMENI	P15325 emericeella

ALIGNMENTS

RESULT 1
MYCT_HUMAN
ID MYCT_HUMAN STANDARD; PRT; 629 AA.
AC Q96QE2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proton myo-inositol co-transporter (Hmit).
GN SLC2A13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21391710; PubMed=11500374;
RA Uldry M., Ibberson M., Horisberger J.-D., Chatton J.-Y.,
RA Riederer B.M., Thorens B.;
RT "Identification of a mammalian H(+)-myo-inositol symporter expressed
predominantly in the brain.";
RL EMBO J. 20:4467-4477(2001).
CC -!- FUNCTION: H(+)-myo-inositol co-transporter. Can also transport
related stereoisomers.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Predominantly expressed in the brain.
CC -!- PTM: Glycosylated.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; AJ315644; CAC51116.1; -.
Genew; HGNC:15956; SLC2A13.
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub_transporter.
InterPro; IPR005829; Sug_transporter.
InterPro; IPR003663; Sugar_transpt.
Pfam; PF00083; sugar_tr; 1.
PRINTS; PR00171; SUGTRNSPORT.
TIGRFAMs; TIGR00879; SP; 1.
PROSITE; PS50850; MFS; 1.
PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
Transmembrane; Sugar_transport; Glycoprotein.
DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
TRANSMEM 58 78 1 (POTENTIAL).
DOMAIN 79 106 EXTRACELLULAR (POTENTIAL).
TRANSMEM 107 127 2 (POTENTIAL).
DOMAIN 128 129 CYTOPLASMIC (POTENTIAL).
TRANSMEM 130 150 3 (POTENTIAL).
DOMAIN 151 159 EXTRACELLULAR (POTENTIAL).

FT	TRANSMEM	160	180	4 (POTENTIAL).	
FT	DOMAIN	181	193	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	194	214	5 (POTENTIAL).	
FT	DOMAIN	215	220	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	221	241	6 (POTENTIAL).	
FT	DOMAIN	242	305	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	306	326	7 (POTENTIAL).	
FT	DOMAIN	327	344	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	345	365	8 (POTENTIAL).	
FT	DOMAIN	366	374	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	375	395	9 (POTENTIAL).	
FT	DOMAIN	396	489	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	490	510	10 (POTENTIAL).	
FT	DOMAIN	511	530	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	531	551	11 (POTENTIAL).	
FT	DOMAIN	552	554	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	555	575	12 (POTENTIAL).	
FT	DOMAIN	576	629	CYTOPLASMIC (POTENTIAL).	
FT	CARBOHYD	414	414	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	439	439	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	466	466	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SQ	SEQUENCE	629 AA; 68203 MW; 8AD79C56C790814E CRC64;			
Query Match 25.1%; Score 642.5; DB 1; Length 629;					
Best Local Similarity 28.9%; Pred. No. 2e-35;					
Matches 158; Conservative 88; Mismatches 205; Indels 95; Gaps 8;					
QY	20	YASICAILASMASVILGIDIGVMSGAAMYIKDLNITDVQLEILIGILSLYSLFGSFAGA	79		
Db	61	FVYVAVFSALGGFLGYDTGVSGAMLLKRLQSLDALWQELLVSVSTVGAAVSAALAGG	120		
QY	80	RTSDRIGRRLTVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGYGGMIAPVYTAEI	139		
Db	121	ALNGVFGRRRAAILLASALFTAGSAVLAANNNKETLLAGRLVVLGIGIGIASMTVPVYIAEV	180		
QY	140	SPAASRGFLTTPPEVFINGILLGYSNFAFARLPDLHGWVRVLAIGAVPSGLLALLVFC	199		
Db	181	SPNLRGRRLVTINTLFTITGGOFFASVVDGAFSYLQKD-GWRYMLGLAXVPAVIOFFGFLF	239		
QY	200	MPESPRWLVKGRADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGVVTVPGKEQ	259		
Db	240	LPESPRWLIQGTQKARRILSQMRGN-QTIDEYDSIK-----NNIEEEKEV	287		
QY	260	GGGELQVWKKLILSPTPAVRRIILSAVGLHFFQASGSDSVVQYSARLFXSAGITDDNKL	319		
Db	288	GSAGPVICRMLSYPT---RRALIVGCGLQMFQQLSGINTIMYYSATILQMSGVEDDRLA	344		
QY	320	LGVTCAVGVTKTFFILVATFLLDRAGRRPRLLLISTGGMVSLICLGSGLTVAGH-----	373		
Db	345	IWLASVTAFTNFIFTLVGVWLVKVGRRKLTFGSLAGTVALIILALGFVLSAQVSPRIT	404		
QY	374	-----HPD-----	376		
Db	405	FKPIAPSGQNATCTRYSYNCEMCLDPDCGFCXKMNKSTVIDSSCVPVNKASTNEAAWGRC	464		
QY	377	---TK-----VAVAVALC-----IASTLSYIAFFSIGLGPITGVYTSEIFPLQVRAL	420		
Db	465	ENETKFKTEDIFWAYNFCPTPYSWTALLGLILYVFPAGMGMPWTVNSEIYPLWARST	524		
QY	421	GFAVGVASNEVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVWVFFFTCLPETRGRITLE	480		
Db	525	GNACSSGINWIFNLVSLTFLHTAEVLTYYGAFFLYAGFAAVGLLFIYGCCLPETHGKLE	584		
QY	481	EMGKLF 486			
Db	585	EIESLF 590			
RESULT 2					
CSBC BACSU					
ID	CSBC BACSU STANDARD; PRT; 461 AA.				
AC	P46333; O32289;				
DT	01-NOV-1995 (Rel. 32, Created)				

DT	10-OCT-2003 (Rel. 42, Last sequence update)	
DT	10-OCT-2003 (Rel. 42, Last annotation update)	
DE	Probable metabolite transport protein csbc.	
GN	CSBC OR SS92BR OR BSU39810.	
OS	Bacillus subtilis.	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
OX	NCBI_TaxID=1423;	
RN	[1] SEQUENCE FROM N.A.	
RP	STRAIN=168 / BGSC1A1;	
RC	MEDLINE=96093926; PubMed=7584049;	
RX	Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;	
RA	"Cloning and sequencing of a 36-kb region of the Bacillus subtilis	
RT	genome between the gnt and iol operons.";	
RL	DNA Res. 2:61-69(1995).	
RN	[2] REVISIONS.	
RP	Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;	
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.	
RN	[3] SEQUENCE FROM N.A.	
RP	STRAIN=168;	
RC	MEDLINE=98044033; PubMed=9384377;	
RX	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,	
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,	
RA	Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,	
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,	
RA	Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,	
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,	
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,	
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,	
RA	Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,	
RA	Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,	
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,	
RA	Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,	
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,	
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,	
RA	Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,	
RA	Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,	
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,	
RA	Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,	
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,	
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,	
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,	
RA	Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,	
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,	
RA	Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,	
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,	
RA	Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,	
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,	
RA	Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;	
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus	
RL	subtilis.";	
RT	Nature 390:249-256(1997).	
RN	[4] CHARACTERIZATION.	
RP	MEDLINE=99303315; PubMed=10376822;	
RX	Akbar S., Lee S.Y., Boylan S.A., Price C.W.;	
RA	"Two genes from Bacillus subtilis under the sole control of the	
RT	general stress transcription factor sigmaB.";	
RL	Microbiology 145:1069-1078(1999).	
CC	-!- FUNCTION: Could serve either a nutritional or an osmotic	
CC	protection function.	
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).	
CC	-!- SIMILARITY: Belongs to the sugar transporter family.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	

FT	TRANSMEM	402	422	12 (POTENTIAL).	
FT	DOMAIN	423	457	CYTOPLASMIC (POTENTIAL).	
SQ	SEQUENCE	457 AA;	49199 MW;	E097EB2B67A92F67 CRC64;	
Query Match 23.1%; Score 591.5; DB 1; Length 457;					
Best Local Similarity 30.6%; Pred. No. 3.4e-32;					
Matches 151; Conservative 101; Mismatches 184; Indels 57; Gaps 11;					
QY	29	SMASVILGYDIGVMSGAAMYIKKDLNITDVQLEILIGILSLYSLFGSFACARTSDRIGRR	88		
Db	14	ALGGLLFGYDTGVISGAILFIQKQMNGLSWQQGWVVSAILGAILGAILGPSSDRFGR	73		
QY	89	LTVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGYGGMIAPVYTAETISPAASRGFL	148		
Db	74	KLLLSAIIFFVGALGSAFSPFWTLIISRIILGMVGAASALIPTYLAEAPSDKRGTV	133		
QY	149	TFPEVFINIGILLGYLSNFPAPARPLHLGWRVMAIGAIVPSGLLALLVFCMPESPRWL	208		
Db	134	SSLQQLMVTGILLAYITNYSFS--GFYTGWRWMLGFAAIPAALLFLGGLLPESPRFL	191		
QY	209	LKGRLLADARAVLEKTSATPEEAER-LADIKAAAGIPKGLDGDVTVVPCKEQGGGELQV	267		
Db	192	KSGHLDEARHVLDTNKHQDVAVNKEINDIQESA-----KIVSGGWSELF	236		
QY	268	KKLILSPTPAVRRILLSAVGLHFFQQASGSDSVVQYSARLFKSAGITDDNKLGLVTCA--	325		
Db	237	GKM-----VRPSLIIGIGLAIFQQVMGCNTLVYAPTIFTDVG-----FGVSAALL	282		
QY	326	----VGVTKTFILVATFLDRAGRRPLLLISTGMIVSLICLGSGLTVAGHHPTKVAW	381		
Db	283	AHIGIGFNVIVTAIAVAIMDKIDRKIVNIGAVGSGISLFVMSIGMKFSGSQTAIIIS	342		
QY	382	AVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFVAVGASNRVTSAVISMTFL	441		
Db	343	VIALTV----YIAFFSATWGPVWVMWIGEVFPLNIRGLNSFASVINWTANMIVSLTFP	397		
QY	442	SLSKAITIGSGFFLYSGIAAVAVVFFF-TCLPETRGRTLEEMGKLFMPDTCMAEEAEDA	500		
Db	398	SLLDFFGT-GSLFIGYGLCFASIFVQKVFETRNRSLD-----IEATLRAKTGEDA	450		
QY	501	AAKEKVELPSSK	513		
Db	451	A-----ELSTTK	457		
RESULT 4					
ARAE	ECOLI	STANDARD;	PRT;	472 AA.	
ID	ARAE	ECOLI			
AC	P09830;	Q46937;			
DT	01-MAR-1989	(Rel. 10, Created)			
DT	01-MAR-1989	(Rel. 10, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	Arabinose-proton symporter (Arabinose transporter).				
GN	ARAE	OR B2841	OR Z4161	OR ECS3698.	
OS	Escherichia coli,	and			
OS	Escherichia coli	O157:H7.			
OC	Bacteria;	Proteobacteria;	Gammaproteobacteria;	Enterobacteriales;	
OC	Enterobacteriaceae;	Escherichia.			
OX	NCBI	TaxID=562,	83334;		
RN	[1]				
RP	SEQUENCE	FROM N.A.			
RX	MEDLINE=87115869;	PubMed=3543693;			
RA	Maiden M.C.J.,	Davis E.O.,	Baldwin S.A.,	Moore D.C.M.,	
RA	Henderson P.J.F.;				
RT	"Mammalian and bacterial sugar transport proteins are homologous."				
RL	Nature	325:641-643	(1987).		
RN	[2]				
RP	SEQUENCE	FROM N.A.			
RC	STRAIN=K12 /	JM2433;			
RX	MEDLINE=88228015;	PubMed=2836407;			
RA	Maiden M.C.J.,	Jones-Mortimer M.C.,	Henderson P.J.F.;		
RT	"The cloning, DNA sequence, and overexpression of the gene araE				
RT	coding for arabinose-proton symport in Escherichia coli K12."				
RL	J. Biol. Chem.	263:8003-8010	(1988).		
RN	[3]				
RP	SEQUENCE	FROM N.A.			
RC	STRAIN=K12 /	MG1655;			
RX	MEDLINE=97426617;	PubMed=9278503;			
RA	Blattner F.R.,	Plunkett G. III,	Bloch C.A.,	Perna N.T.,	Burland V.,
RA	Riley M.,	Collado-Vides J.,	Glasner J.D.,	Rode C.K.,	Mayhew G.F.,
RA	Gregor J.,	Davis N.W.,	Kirkpatrick H.A.,	Goeden M.A.,	Rose D.J.,
RA	Mau B.,	Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12."				
RL	Science	277:1453-1474	(1997).		
RN	[4]				
RP	SEQUENCE	FROM N.A.			
RC	STRAIN=O157:H7 /	EDL933 /	ATCC 700927;		
RX	MEDLINE=21074935;	PubMed=11206551;			
RA	Perna N.T.,	Plunkett G. III,	Burland V.,	Mau B.,	Glasner J.D.,
RA	Rose D.J.,	Mayhew G.F.,	Evans P.S.,	Gregor J.,	Kirkpatrick H.A.,
RA	Posfai G.,	Hackett J.,	Klink S.,	Boutin A.,	Shao Y.,
RA	Grotbeck E.J.,	Davis N.W.,	Lim A.,	Dimalanta E.T.,	Potamousis K.,
RA	Apodaca J.,	Anantharaman T.S.,	Lin J.,	Yen G.,	Schwartz D.C.,
RA	Welch R.A.,	Blattner F.R.;			
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."				
RL	Nature	409:529-533	(2001).		
RN	[5]				
RP	SEQUENCE	FROM N.A.			
RC	STRAIN=O157:H7 /	RIMD 0509952;			
RX	MEDLINE=21156231;	PubMed=11258796;			
RA	Hayashi T.,	Makino K.,	Ohnishi M.,	Kurokawa K.,	Ishii K.,
RA	Han C.-G.,	Ohtsubo E.,	Nakayama K.,	Murata T.,	Tanaka M.,
RA	Iida T.,	Takami H.,	Honda T.,	Sasakawa C.,	Ogasawara N.,
RA	Kuhara S.,	Shiba T.,	Hattori M.,	Shinagawa H.;	
RT	"Complete genome sequence of enterohemorrhagic Escherichia coli				
RT	O157:H7 and genomic comparison with a laboratory strain K-12."				
RL	DNA Res.	8:11-22	(2001).		
RN	[6]				
RP	PRELIMINARY	SEQUENCE OF 1-28	FROM N.A.		
RX	MEDLINE=84114868;	PubMed=6319708;			
RA	Stoner C.,	Schleif R.F.;			
RT	"The araE low affinity L-arabinose transport promoter. Cloning,				
RT	sequence, transcription start site and DNA binding sites of				
RT	regulatory proteins."				
RL	J. Mol. Biol.	171:369-381	(1983).		
CC	-!	FUNCTION: Uptake of arabinose across the boundary membrane with			
CC	-!	the concomitant export of a proton (symport system).			
CC	-!	SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.			
CC	-!	SIMILARITY: Belongs to the sugar transporter family.			
CC					
CC	This	SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between	the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the	European Bioinformatics Institute. There are no restrictions on its			
CC	use	by non-profit institutions as long as its content is in no way			
CC	modified	and this statement is not removed. Usage by and for commercial			
CC	entities	requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or	send an email to license@isb-sib.ch).			
CC					
CC	EMBL;	J03732;	AAA23469.1;	-.	
DR	EMBL;	X00272;	CAA25075.1;	-.	
DR	EMBL;	U29581;	AAB40488.1;	-.	
DR	EMBL;	AE000368;	AAC75880.1;	-.	
DR	EMBL;	AE005513;	AAG57953.1;	-.	
DR	EMBL;	AP002563;	BAB37121.1;	-.	
DR	PIR;	B26430;	B26430.		
DR	PIR;	B91091;	B91091.		
DR	PIR;	E85936;	E85936.		
DR	EcoGene;	EG10056;	araE.		
DR	InterPro;	IPR007114;	MFS.		
DR	InterPro;	IPR005828;	Sub_transporter.		
DR	InterPro;	IPR005829;	Sug_transporter.		
DR	InterPro;	IPR003663;	Sugar_transpt.		
DR	Pfam;	PF00083;	sugar tr;	1.	
DR	PRINTS;	PRO0171;	SUGTRNSPORT.		
DR	TIGRFAMs;	TIGR00879;	SP;	1.	
DR	PROSITE;	PS50850;	MFS;	1.	

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT DOMAIN 1 29 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 30 50 1 (POTENTIAL).
FT DOMAIN 51 63 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 64 84 2 (POTENTIAL).
FT DOMAIN 85 91 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 92 112 3 (POTENTIAL).
FT DOMAIN 113 114 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 115 135 4 (POTENTIAL).
FT DOMAIN 136 154 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 5 (POTENTIAL).
FT DOMAIN 176 178 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 179 199 6 (POTENTIAL).
FT DOMAIN 200 257 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 258 278 7 (POTENTIAL).
FT DOMAIN 279 297 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 298 318 8 (POTENTIAL).
FT DOMAIN 319 325 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 326 346 9 (POTENTIAL).
FT DOMAIN 347 361 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 362 382 10 (POTENTIAL).
FT DOMAIN 383 404 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 405 425 11 (POTENTIAL).
FT DOMAIN 426 427 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 428 446 12 (POTENTIAL).
FT DOMAIN 447 472 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 26 28 SVA -> YDR (IN REF. 6).
SQ SEQUENCE 472 AA; 51684 MW; 411990A441D44393 CRC64;

Query Match 22.9%; Score 586.5; DB 1; Length 472;
Best Local Similarity 29.9%; Pred. No. 7.5e-32;
Matches 147; Conservative 104; Mismatches 197; Indels 43; Gaps 9;

QY 1 MASDELAKAVEPRKKGNVKXASI-CAILASMASVILGYDIGVMSGAAMYIKKDLNITDVQ 59
Db 1 MVTINTESALTPRSLRDRTRMMNFVSAAVAGLFLGDIGVIAGALPFITDHFVLTSRL 60

QY 60 LEILIGILSLYSLFGSFAGARTSDRIGRRLLTVFAAVIFFVGSLLMGFAVNYGMLMAGR 119
Db 61 QEWVSSMMLGAAIGALFNGLWLSFRLGRKYSMLMAGAILFVLGSLGSAFATSVEMLIAARV 120

QY 120 VAGVGVGYGGMIAPVYTABIISPAASRGFLTTFPEVFINIGILLGLYSNFAFARLPHLG- 178
Db 121 VLGIAGVGIASYTAPLYLSMASENVRGKMISMVQLMVTGLIVLAFLSDTAFS---YSGN 176

QY 179 WRVMLAIGAVPSGLLALVFCMPESPRLWLKGRLLADARAVLEKTSATPEBAERLADIK 238
Db 177 WRAMLGVLALPAVLLIILVVFLPNSPRWLAEKGRHIEAEVLRMLRDTSEKAREELNEIR 236

QY 239 AAAGIPKGLDGDVVTVPGKEQGGELQVWKKLILSPTPAVRRILLSAVGLHFFQQAQSGSD 298
Db 237 ESLKLJ-----KQGG-----WALPKINRN--VRRAVFLGMLLQAMQOFTGMN 275

QY 299 SVVQYSARLFKSAGITDDNKLGVTCVAVGVTKTFFILVATFLLDRAGRRPLLIIISTGMI 358
Db 276 IIMYYAPRIFKMGAGFTTTEQQMIATLVVGLTFMEATFIAVFTVDKAGRKPKLIGFSVMA 335

QY 359 VSLICLG-----SGLTVAGHHPDTKVAVAWAVALCIASTLSYIAFFSIGLGPITGVYTSE 411
Db 336 LGTLVLGYCLMQFDNGTASSG-----LSW---LSVGMTMMCAGYAMSAPVWILCSE 386

QY 412 IFPLQVRALGFVGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAVAVWVFFFTCL 471
Db 387 IQPLKCRDFGITCSTTTNWVSNMIIGATFTLLDSIGAAGTFWLTYALTALNIAFVGITFWLI 446

QY 472 PETRGRTLEEM 482
Db 447 PETKNVTLEHI 457

RESULT 5
ITR2_SCHPO
ID ITR2_SCHPO STANDARD; PRT; 557 AA.
AC P87110; P78901;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 2.
GN ITR2 OR SPAC20G8.03.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 h90;
RX MEDLINE=98228265; PubMed=9560432;
RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
RA Rusu M., Poitelea M., Edenharter L., Schweingruber M.E.;
RT "Exogenous inositol and genes responsible for inositol transport are
RT required for mating and sporulation in Shizosaccharomyces pombe.";
RL Curr. Genet. 33:255-261(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Harris P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA Holroyd S., Hornsby T., Jones M., Leather S., McDonald S., McLean J.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [3]
RP SEQUENCE OF 166-541 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs.";
RL DNA Res. 4:363-369(1997).
CC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC

FT	TRANSMEM	179	199	6 (POTENTIAL).
FT	DOMAIN	200	257	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	258	278	7 (POTENTIAL).
FT	DOMAIN	279	297	PERIPLASMIC (POTENTIAL).
FT	TRANSMEM	298	318	8 (POTENTIAL).
FT	DOMAIN	319	325	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	326	346	9 (POTENTIAL).
FT	DOMAIN	347	361	PERIPLASMIC (POTENTIAL).
FT	TRANSMEM	362	382	10 (POTENTIAL).
FT	DOMAIN	383	404	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	405	425	11 (POTENTIAL).
FT	DOMAIN	426	427	PERIPLASMIC (POTENTIAL).
FT	TRANSMEM	428	446	12 (POTENTIAL).
FT	DOMAIN	447	472	CYTOPLASMIC (POTENTIAL).
SQ	SEQUENCE	472 AA;	51732 MW;	410021E1BEE3D96E CRC64;
Query Match				
Best Local Similarity 22.6%; Score 579; DB 1; Length 472;				
Matches 142; Conservativity 30.5%; Pred. No. 2.4e-31;				
Matches 142; Conservativity 96; Mismatches 186; Indels 42; Gaps 8;				
QY	25	ATLASMASVILGYDIGVMSGAAMYIKKDLNITDVQLEILIGILSLYSLFGSFAGARTSDR	84	
Db	26	SIAAAVAGLLFGLDIGVIAGALPFITDHFVLSRLQEWVSSMMLGAAIGALFNGWLSFR	85	
QY	85	IGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGYGMIAPVYTAEISPAAS	144	
Db	86	LGRKYSLMVGAVLFVAGSVGSFAFATSVEMLLVARIVLGVAVGIASTYAPLYLSEMAENV	145	
QY	145	RGFLTTFPEVFINIGILLGVLNSFAFARLPHLG-WRVMLAIGAVPSGLLALLVFCMPES	203	
Db	146	RGMWISMYQLMVTGLGVNVAFLSDTAFS----YSGNWRAMGLVLPVAVLLIILVIFLPNS	201	
QY	204	PRWLVLKGRLDARAVLEKTSATPEEAAERLADIAAAGIPKGLDGDVVTVPGKEQGGGE	263	
Db	202	PRWLAEKGRHVEAEVLRMLRDTSEKARDELNEIRESKL-----KQGG--	245	
QY	264	LQVWKKLILSPTPAVRRIILSAVGLHFFQAGSDSVVQYSARLFKSAGITDDNKLLGVT	323	
Db	246	---W--ALEKVNRRVRAVFLGMLLQAMQQTGMNIIMYAPRIFKMGFTTTEQQMVAT	300	
QY	324	CAVGVTKTFFILVATFLDRAGRRPLLLISTGMIVSLICLG-----SGLTVAGHHPD	376	
Db	301	LVVGLTFMEATFIAVFTVDKAGRKPAKIGFSVMAIGTLVLGYCLMQFDNGTASSG----	356	
QY	377	TKVAVAVALCIASLTSLYIAFFSIGLPIITGVYTSBIFPLQVRALGFAVGVASNRVTSAVI	436	
Db	357	--LSW---LSVGMTMMCIAGYAMSAAPVVWILCSEIQPLKCRDFGITCSTTNVWSNMII	411	
QY	437	SMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGRITLEEM	482	
Db	412	GATFLTLLDAIGAAGTFWLYTALNVAFIGVTFWLPETKNVTLEHI	457	

RESULT 7

ID	YFIG_BACSU	STANDARD;	PRT;	482 AA.
AC	P54723;			
DT	01-OCT-1996	(Rel. 34, Created)		
DT	01-OCT-1996	(Rel. 34, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Hypothetical metabolite transport protein yfiG.			
GN	YFIG OR BSU08260.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=96262713;	PubMed=8704981;		
RA	Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;			
RT	"Determination of a 12 kb nucleotide sequence around the 76 degrees region of the Bacillus subtilis chromosome."			
RL	Microbiology 142:1417-1421(1996).			

RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=98044033;	PubMed=9384377;		
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Borchert S., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Brignell S.C., Bron S., Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Carter N.M., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;			
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus subtilis";			
RL	Nature 390:249-256(1997).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).			
CC	-!- SIMILARITY: Belongs to the sugar transporter family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL;	D50543;	BAA09111.1;	--
DR	EMBL;	Z99108;	CAB12655.1;	--
DR	PIR;	B69803;	B69803.	
DR	Subtilist;	BGI1854;	yfiG.	
DR	InterPro;	IPR007114;	MFS.	
DR	InterPro;	IPR005828;	Sub_transporter.	
DR	InterPro;	IPR005829;	Sug_transporter.	
DR	InterPro;	IPR003663;	Sugar_transpt.	
DR	Pfam;	PF00083;	sugar_tr; 1.	
DR	PRINTS;	PR00171;	SUGTRNSPORT.	
DR	TIGRFAMs;	TIGR00879;	SP; 1.	
DR	PROSITE;	PS50850;	MFS; 1.	
DR	PROSITE;	PS00216;	SUGAR_TRANSPORT_1; FALSE_NEG.	
DR	PROSITE;	PS00217;	SUGAR_TRANSPORT_2; 1.	
KW	Hypothetical protein; Transport; Transmembrane; Complete proteome.			
FT	DOMAIN	1	29	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	30	50	1 (POTENTIAL).
FT	DOMAIN	51	59	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	60	80	2 (POTENTIAL).
FT	DOMAIN	81	92	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	93	113	3 (POTENTIAL).
FT	DOMAIN	114	120	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	121	141	4 (POTENTIAL).
FT	DOMAIN	142	155	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	156	176	5 (POTENTIAL).
FT	DOMAIN	177	184	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	185	205	6 (POTENTIAL).

FT DOMAIN 206 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 284 7 (POTENTIAL).
FT DOMAIN 285 301 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 302 322 8 (POTENTIAL).
FT DOMAIN 323 331 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 332 352 9 (POTENTIAL).
FT TRANSMEM 353 373 10 (POTENTIAL).
FT DOMAIN 374 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 11 (POTENTIAL).
FT DOMAIN 422 444 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 445 482 12 (POTENTIAL).
SQ SEQUENCE 482 AA; 52756 MW; 90851C4F4C48EE01 CRC64;

Query Match 22.2%; Score 568.5; DB 1; Length 482;
Best Local Similarity 31.1%; Pred. No. 1.2e-30;
Matches 146; Conservative 93; Mismatches 191; Indels 39; Gaps 10;

Qy 26 ILASMASVILGYDIGVMSGAAMYIKK--DLNITDVQLEILIGILSLYSLFGSFAGARTSD 83
Db 26 LVSTFGGLLFYDGTGVINGALPFMATAGQLNLTPTVTEGLVASSLLGAAFGAMFGRLSD 85
Qy 84 RIGRRLTVFAAVIFFVGSLLMGFAVNYGMLNAGRPVAGVGVGYGGMIAPIVTAETISPA 143
Db 86 RHGRKTIYLLALLFIAATLCTFSPNASVMIAFRFLLGLAVGCASVTPTFLAEISPAE 145
Qy 144 SRGFLTTFPEVFINIGILLYLSNFAFARLPHLG-----WRVMLAIGAVSPSGLLALLVF 198
Db 146 RRGRIVTQNELMIVIGQLLAYTFN--AIIGSTMGESANWRYMLVIAITLPAVVLWFGML 202
Qy 199 CMPESPRWLVLKGRADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGDVVTVPKGE 258
Db 203 IVPESPRWLAAGRMGDALRVLRQI-REDSQAQBEIKEIKHA-----IEG-----TAKK 250
Qy 259 QGGGELQVWKKLILSPTPAVRILLSSAVGLHFFQOAGSDSVVQYSARLFSAGITDDNK 318
Db 251 AGPHDFQ-----EPWIRRLIFIGIAIVQITGVNSIMYYGTETILREAGFQTEAA 301
Qy 319 LLGVTCAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLICLGSGLTVAGHHPDTK 378
Db 302 LIG-NIANGVISVIAVIFGFWLLGKVRRRPMLTIQIGTMTALLIGILSIVLEGTP--- 357
Qy 379 VAWAVALCIAS-TLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFAVGASNRVTSAVIS 437
Db 358 ---ALPYVVLSTILFLAFQQTASTVTWMLSEIFPMHVRGLGMGISTFCLMTANFLIG 414
Qy 438 MTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPRTGRTELEEMKLF 486
Db 415 FTFPILLNHIGMSATFFIFVAVNLAIFLVKKYVPETKGRSLEQLEHSF 463

RESULT 8
ITR1_YEAST
ID ITR1_YEAST STANDARD; PRT; 584 AA.
AC P30605;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 1.
GN ITR1 OR YDR497C OR D9719.3.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91250431; PubMed=2040626;
RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT "Isolation and characterization of two distinct myo-inositol
transporter genes of Saccharomyces cerevisiae."
RL J. Biol. Chem. 266:11184-11191(1991).
RN [2]
RP SEQUENCE FROM N.A.

RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
RA Berno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Winant A., Yelton M., Botstein D., Davis R.W.,
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAJOR TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90352; BAA14366.1; -.
CC EMBL; U33057; AAB64939.1; -.
CC PIR; S69555; S69555.
CC Geronline; 140989; -.
CC SGD; S0002905; ITR1.
CC GO; GO:0005365; F:myo-inositol transporter activity; IMP.
CC GO; GO:0015798; P:myo-inositol transport; IMP.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PRO0171; SUGTRNSPORT.
CC TIGRFAMS; TIGR00879; SP; 1.
CC PROSITE; PS50850; MFS; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar_transport; Glycoprotein.
FT DOMAIN 1 81 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 82 102 1 (POTENTIAL).
FT DOMAIN 103 129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 130 150 2 (POTENTIAL).
FT DOMAIN 151 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 184 3 (POTENTIAL).
FT DOMAIN 185 186 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 187 207 4 (POTENTIAL).
FT DOMAIN 208 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 5 (POTENTIAL).
FT DOMAIN 237 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 267 6 (POTENTIAL).
FT DOMAIN 268 349 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 350 370 7 (POTENTIAL).
FT DOMAIN 371 376 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 377 397 8 (POTENTIAL).
FT DOMAIN 398 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 9 (POTENTIAL).
FT DOMAIN 422 441 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 442 462 10 (POTENTIAL).
FT DOMAIN 463 486 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 487 507 11 (POTENTIAL).
FT DOMAIN 508 510 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 511 531 12 (POTENTIAL).
FT DOMAIN 532 584 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 371 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 43 TL -> HI (IN REF. 1).
SQ SEQUENCE 584 AA; 63569 MW; 42543E30A102DC65 CRC64;

Query Match 22.1%; Score 565; DB 1; Length 584;
Best Local Similarity 29.5%; Pred. No. 2.5e-30;
Matches 150; Conservative 109; Mismatches 183; Indels 66; Gaps 13;
Qy 27 LASMASVILGYDIGVMSGAAMYIKKDLN---ITDVQLEILIGILSLYSLFGSFAGARTSD 83
Db 92 VASISGFMFGYDGTGYISSALISIGTDLDHKULTYGEKEIVTAATSLGALITSIFAGTAAD 151

QY 84 RIGRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGMIAPIVYTAIESPAA 143
Db 152 IFGRKRCLMGSLMFMFVIGAILQVSAHTFQWAVGRLLMGFGVIGSLIAPLFISEIAPKM 211
QY 144 SRGFLTTFFPEVINIGIL-----LGVLNFAFARLPLHLGWRVLAIGAVPSGLLAL 195
Db 212 IRGLRTVINSLWLTGGQLVAYGCGAGLNVNN-----GWRILVGLSLIPTAVQFT 261
QY 196 LVFCMPESPRLVLKGRLLADARAVLEKT-SATPEEAAERLADIKAAAGIPKGLDGVVT- 253
Db 262 CLCFLPDTPRYYVMKGDRLARATEVLKRSYTDISEIIEKVE-----ELVTL 308
QY 254 ---VPGKEQGGELQVWKKL-ILSPTPAVRRLILLSAVGLHFFQOAGSDSVVQYSARLFK 309
Db 309 NQSIPIGKNV---PEKVNTIKELHTVPSNLRALIIGCGLQAIQQFTGWSNLMYFSGTIPE 365
QY 310 SAGITDDNKLGLVTCVAVGVTKTFILVATFLDDRARRPLLLISTGGMIVSLICLGSGLT 369
Db 366 TVGFKNSS---AVSIIVSGTNFIFTLVAFPSIDKIGRRRTILLIGLPGMTALVV---CS 418
QY 370 VAGHHPDTKVAVAVALCIASLTLS-----YIAFFSICLGPITGVYTSEIFPLQV 417
Db 419 IAFHFLGIKFDGAVAVVSSGFSWGIVIIIVFAAFYALGIGTVPW-QQSELFQNV 477
QY 418 RALGFAVGASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAVAVVFFFTCLPETRGR 477
Db 478 RGIGTSYATATNWAGSLVIASTFLTMLQNITPAGTFAFFAGLSCLSTIFCYFCYPELSGL 537
QY 478 TLEEMGKLFMPDPTGMAEEAEADAKEK 505
Db 538 ELEEVQTL---KDGFNKASKALAKKR 562
RESULT 9
ITR2 YEAST
ID ITR2 YEAST STANDARD; PRT; 612 AA.
AC P30606;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 2.
GN ITR2 OR YOL103W OR HRB612.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=91250431; PubMed=2040626;
RA Nikawa J.-I., Tsukagoshi Y., Yamashita S.;
RT "Isolation and characterization of two distinct myo-inositol
transporter genes of Saccharomyces cerevisiae.";
RL J. Biol. Chem. 266:11184-11191(1991).
RN [2]
SEQUENCE FROM N.A.
MEDLINE=96076631; PubMed=7502582;
RA Vandenbol M., Durand P., Portetelle D., Hilger F.;
RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV
including the Tyl-H3 retrotransposon, the sufl(+) frameshift
suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-ia and a
delta element.";
RL Yeast 11:1059-1075(1995).
CC -!- FUNCTION: MINOR TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC EMBL; D90353; BAA14367.1; -
CC EMBL; Z48149; CAA88159.1; -
DR EMBL; Z74845; CAA99119.1; -
DR PIR; B40538; B40538.
DR Germline; 143525; -
DR SGD; S0005463; ITR2.
DR GO; GO:0005365; F:myo-inositol transporter activity; IMP.
DR GO; GO:0015798; P:myo-inositol transport; IMP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 107 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 108 128 1 (POTENTIAL).
FT DOMAIN 129 155 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 156 176 2 (POTENTIAL).
FT DOMAIN 177 182 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 183 203 3 (POTENTIAL).
FT DOMAIN 204 212 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 213 233 4 (POTENTIAL).
FT DOMAIN 234 241 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 242 262 5 (POTENTIAL).
FT DOMAIN 263 272 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 273 293 6 (POTENTIAL).
FT DOMAIN 294 362 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 363 383 7 (POTENTIAL).
FT DOMAIN 384 402 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 403 423 8 (POTENTIAL).
FT DOMAIN 424 426 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 427 447 9 (POTENTIAL).
FT DOMAIN 448 467 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 468 488 10 (POTENTIAL).
FT DOMAIN 489 512 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 513 533 11 (POTENTIAL).
FT DOMAIN 534 536 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 537 557 12 (POTENTIAL).
FT DOMAIN 558 612 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 612 AA; 67041 MW; 48E733102BAD7298 CRC64;

Query Match 22.0%; Score 563; DB 1; Length 612;
Best Local Similarity 29.3%; Pred. No. 3.6e-30;
Matches 149; Conservative 108; Mismatches 184; Indels 68; Gaps 16;
QY 27 LASMASVILGYDIGVMSGAAMYIKKDLN---ITDVQLLEILGILSLYSLFGSFAGARTSD 83
Db 118 VASISGFMFGYDTGYISSALISINRDLNKKVLTGYGEKELITAATSLGALITSVGAGTAAD 177
QY 84 RIGRLTVVFAAIVFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGMIAPIVYTAIESPAA 143
Db 178 VFGRRPCLMFSLMFLIGAILQITAHKFWQMAAGRLIMFGVIGSLISPLFISEIAPKM 237
QY 144 SRGFLTTFFPEVINIGILGY-----LSNFAFARLPLHLGWRVLAIGAVPSGLLAL 195
Db 238 IRGLRTVINSLWLTGGQLVAYGCGAGLNVNN-----GWRILVGLSLIPT-VLQF 286
QY 196 LVFC-MPESPRWLVLKGRLLADARAVLEKTSATPE-----EAAERLADIKAAAGIPKGLDG 249
Db 287 SFFCELDPDTPRYVMKGDRLARATEVLKRSYVTEDEIIDQKVEELSSLNQ----- 336
QY 250 DVVTVPKEQGGELQVWKKL-ILSPTPAVRRLILLSAVGLHFFQOAGSDSVVQYSARLF 308
Db 337 ---SIPGKNP---ITKFWNMVKELHTVPSNFRALIIGCGLQAIQQFTGWSNLMYFSGTIF 390


```

QY 309 KSAGITDDNKLGLGVTCAVGVTKTFFFILVATFLDDRGRPLLLISTGGMIVSL-ICL--- 366
DQ 391 ETVGPKSS--AVSIIVSGINPVFTLIAFFCIDKIGRRYILLGLPGMTVALVICAI 447
QY 365 -----GSLGTLVAGHHPTDKVAVAWALCIASTLSYIAFFSIGLGPITGVYTSIFPLQ 416
DQ 448 HFLGIKENGADAVAS--DGFSSWGIVI-IVFIIVYAAPYALGIGTVPW-QQSELPQN 502
QY 417 VRALGFAVGASNRVTSAVISMTFELSLSKAITIGGSFFLYSGIAAVALVFFFTCLPETRG 476
DQ 503 VRGVGTSYATATNWAGSLVIASTFLTMLQNITPTGTFSFAGVACLSTIFCYFCYPBELSG 562
QY 477 RTLEEMGKLFPGMPDTCMAEEAEADAAGEK 505
DQ 563 LELEEVQTIL---KDFGNKASKALAKKR 588

RESULT 10
GALP_ECOLI STANDARD; PRT; 464 AA.
AC P37021;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Galactose-proton symporter (Galactose transporter).
GN GALP OR B2943 OR C3529.
OS Escherichia coli, and
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI TaxID=562, 217992;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA Roberts P.E.;
RL Thesis (1992), University of Cambridge, U.K.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: UPTAKE OF GALACTOSE ACROSS THE BOUNDARY MEMBRANE WITH
CC THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U28377; AAA69110.1; -.
DR EMBL; AE000377; AAC75980.1; -.
DR EMBL; AE016766; AAN81977.1; ALT_INIT.
DR PIR; F65079; F65079.

```


Db 392 ATNWIAMIVGATFLTMLNTIGNANTFWYAAALNVLFILLTLWLVPETKHSVLSLEHIERNL 451
QY 475 -RGRTLEEMG 483
Db 452 MKGRKLEIG 461

RESULT 11

ITR1_SCHPO STANDARD; PRT; 575 AA.
ID ITR1_SCHPO
AC Q10286;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myo-inositol transporter 1.
GN ITR1 OR SPAC4F8.15 OR SPAC7D4.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 h90;
RX MEDLINE=98228265; PubMed=9560432;
RA Niederberger C., Graub R., Schweingruber A.-M., Fankhauser H.,
RA Rusu M., Poitelea M., Edenharter L., Schweingruber M.E.;
RT "Exogenous inositol and genes responsible for inositol transport are
RL required for mating and sporulation in Shizosaccharomyces pombe.";
RN Curr. Genet. 33:255-261(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).

CC -!- FUNCTION: TRANSPORTER FOR MYO-INOSITOL.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the sugar transporter family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X98622; CAA67211.1; -.
DR EMBL; Z98530; CAB11061.1; -.
DR EMBL; Z99532; CAB16718.1; -.
DR PIR; T43400; T43400.
DR GenedB Spombe; SPAC4F8.15; -.
DR GO; GO:0000747; P:conjugation with cellular fusion; ISS.
DR GO; GO:0006629; P:lipid metabolism; ISS.
DR GO; GO:0007165; P:signal transduction; ISS.
DR GO; GO:0007151; P:sporulation (sensu Saccharomycetes); ISS.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 87 107 1 (POTENTIAL).
FT DOMAIN 108 129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 130 150 2 (POTENTIAL).
FT DOMAIN 151 156 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 157 177 3 (POTENTIAL).
FT DOMAIN 178 186 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 187 207 4 (POTENTIAL).
FT DOMAIN 208 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 5 (POTENTIAL).
FT DOMAIN 237 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 267 6 (POTENTIAL).
FT DOMAIN 268 349 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 350 370 7 (POTENTIAL).
FT DOMAIN 371 376 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 377 397 8 (POTENTIAL).
FT DOMAIN 398 400 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 401 421 9 (POTENTIAL).
FT DOMAIN 422 441 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 442 462 10 (POTENTIAL).
FT DOMAIN 463 486 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 487 507 11 (POTENTIAL).
FT DOMAIN 508 510 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 511 531 12 (POTENTIAL).
FT DOMAIN 532 575 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 575 AA; 62757 MW; 3B7C5EFF86CS96AE CRC64;

Query Match 21.3%; Score 546; DB 1; Length 575;
Best Local Similarity 30.8%; Pred. No. 4.5e-29;
Matches 153; Conservative 86; Mismatches 206; Indels 52; Gaps 12;

QY 19 KYASICAILASMASVILGYDIGVMSGAAMYIKKOL---NITDVQLEILIGILSLYSLFGS 75
Db 85 KVVWVLAFAAGIGGLLFGYDTGVISGALVWIGTSLGGHFLTNGGKEFITSATSLGALLGG 144
QY 76 FAGARTSDRIGRRRLTVWFAAVIFPVGSLLMGFVAVNYGMLMAGRFVAGVGVGGMIAPVY 135
Db 145 ILAGALADFFGRKPVIAIASIIIVGSIQVTAHHLWHMIVGRFVIGWGVGSIASLIIPLY 204
QY 136 TAEISPAASRGFLTTFPEVFINIGILLGYLSNFAFARLPLHLGWRVLAIGAVPSGLLAL 195
Db 205 LSEIAPSKIRGLVIIYVLLITAGQVIAYGIDTAFEH--VHNGWRWVMVGLAMVPAAFQLF 262
QY 196 LVFCMPBPRLVLKGLRADARAVLEK--TSATPEEAERLADIKAAAGIPKGLDGDVVT 253
Db 263 ILIWLPEPRLLVKKERSQEAAYNTLARIYPTAHPEYIKTKLYLIQ--EGVRDPFSG--- 316
QY 254 VPGKEGGGELQVWKLI-----LSPTPAVRRILLSAVGLHFFQQAAGSDSVVQYSARLF 308
Db 317 -----SRWQKIVKTKFELYFNPSNFRALILACGLQAMQQLSGFNSLMYFSSITF 365
QY 309 KSAGITDDNKLGLVTCVAVGVTKTFFILVATFLLDAGRPRPLLSTGGMIVSLICLSGL 368

```
Db      366 EVVGF---NNPTATGLIIAATNFVFTIVAFGVIDFFGRRILLLLTVWGMIAALIV---C 418
QY      369 TVAGHH-----PDKVAVAVALCIASTLSYIAFFSIGLGPITGVYVTSIFPLQVR 418
Db      419 AVAFHFLPKDENGNYTSQSNAAWAIIVLI-SMIVVYVASYASGLGNLPW-QQSELFPMMSVR 476
QY      419 ALGFAVGVASNRVTSAVISMTFLSLSKAITIGSFFLYSGIAAVAVWVFFFTCLPETRGRT 478
Db      477 GLGTGMSTAVNWAGNLGIGASFLTLMSEITPTGTFALYGGLCFLGLWGLALFCYPDLTDYT 536
QY      479 LBEEMGKL-----FGMPDT 491
Db      537 IBEIGELLKHGFGVRES 553
```

RESULT 12

```
XYLE-ECOLI
ID_XYLE-ECOLI STANDARD; PRT; 491 AA.
AC P09098;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-xylose-proton symporter (D-xylose transporter).
GN XYLE OR B4031 OR Z5629 OR ECS5014.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88007632; PubMed=2820984;
RA Davis E.O., Henderson P.J.F.;
RT "The cloning and DNA sequence of the gene xyle for xylose-proton
symport in Escherichia coli K12.";
RL J. Biol. Chem. 262:13928-13932(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87115869; PubMed=3543693;
RA Maiden M.C.J., Davis E.O., Baldwin S.A., Moore D.C.M.,
RA Henderson P.J.F.;
RT "Mammalian and bacterial sugar transport proteins are homologous.";
RL Nature 325:641-643(1987).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
region from 89.2 to 92.8 minutes.";
RL Nucleic Acids Res. 21:5408-5417(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / BDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
```

```
RA      Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 1-192 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=88234001; PubMed=2836810;
RA Francoz E., Dassa E.;
RT "3' end of the maleFG operon in E.coli: localization of the
RT transcription termination site.";
RL Nucleic Acids Res. 16:4097-4109(1988).
CC -!- FUNCTION: Uptake of D-xylose across the boundary membrane with the
CC concomitant export of a proton (symport system).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- INDUCTION: By xylose.
CC -!- MISCELLANEOUS: E.COLI HAS TWO D-XYLOSE TRANSPORT SYSTEMS THAT
CC ACCUMULATE SUGAR AGAINST A CONCENTRATION GRADIENT: THE XYLE SYSTEM
CC WHICH UTILIZES THE ELECTROCHEMICAL GRADIENT OF PROTONS AND THAT IS
CC INSENSITIVE TO COLD OSMOTIC SHOCK AND THE XYLE SYSTEM THAT USES A
CC HIGH-ENERGY PHOSPHATE COMPOUND AND IS SENSITIVE TO COLD OSMOTIC
CC SHOCK.
CC -!- SIMILARITY: Belongs to the sugar transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02812; AAA79016.1; -.
DR EMBL; U00006; AAC43125.1; -.
DR EMBL; AE000476; AAC77001.1; -.
DR EMBL; AE005636; AAG59230.1; -.
DR EMBL; AP002568; BAB38437.1; -.
DR EMBL; X06663; CAA29863.1; -.
DR PIR; A26430; A26430.
DR PIR; B86096; B86096.
DR PIR; F91255; F91255.
DR EcoGene; EG11076; xyle.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT DOMAIN 1 9
FT TRANSMEM 10 30
FT DOMAIN 31 55
FT TRANSMEM 56 76
FT DOMAIN 77 89
FT TRANSMEM 90 110
FT DOMAIN 111 133
FT TRANSMEM 134 154
FT DOMAIN 155 165
FT TRANSMEM 166 186
FT DOMAIN 187 200
FT TRANSMEM 201 221
FT DOMAIN 222 272
FT TRANSMEM 273 293
FT DOMAIN 294 312
FT TRANSMEM 313 333
FT DOMAIN 334 343
FT TRANSMEM 344 364
FT DOMAIN 365 369
FT CYTOPLASMIC (POTENTIAL).
FT 1 (POTENTIAL).
FT PERIPLASMIC (POTENTIAL).
FT 2 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 3 (POTENTIAL).
FT PERIPLASMIC (POTENTIAL).
FT 4 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 5 (POTENTIAL).
FT PERIPLASMIC (POTENTIAL).
FT 6 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 7 (POTENTIAL).
FT PERIPLASMIC (POTENTIAL).
FT 8 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 9 (POTENTIAL).
FT PERIPLASMIC (POTENTIAL).
```

FT	TRANSMEM	370	390	10 (POTENTIAL).
FT	DOMAIN	391	407	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	408	428	11 (POTENTIAL).
FT	DOMAIN	429	442	PERIPLASMIC (POTENTIAL).
FT	TRANSMEM	443	463	12 (POTENTIAL).
FT	DOMAIN	464	491	CYTOPLASMIC (POTENTIAL).
FT	CONFLICT	64	64	A -> V (IN REF. 6).
SQ	SEQUENCE	491 AA;	53608 MW;	2AF1AF9756C0B722 CRC64;
Query Match				
Best Local Similarity 21.3%; Score 544; DB 1; Length 491;				
Matches 158; Conservativity 93; Mismatches 183; Indels 94; Gaps 18;				
Qy	17	NVKYASICAILASMASVILGYDIGVMSGAAMYIK----	KDLNITDVQLEILIGILSLYSL	72
Dd	6	NSSYFISITIVATLGLLFGYDTAVISGTVESLNTVFVAPQNLSESAANSLLCFVASAL	65	
Qy	73	FGSFAGAR----	TSDRIGRRLTVFAAVIFFVGS-----LMGF-AVN-----YGMLMAG-	117
Dd	66	IGCIIGALGGYCSNRFGRDRSLKIAAVLFFISGVGSAWPELGFTSINPDNTVPVYLGY	125	
Qy	118	-----RFVAGVGVGGMIAPVYTAEISPAASRGILTTTPPEVFINIGILLGVLNSPFAF	170	
Dd	126	VPEFVIYRIIGGIVGLASMLSPMYIAELAPAHIRGLVSNQFAIIFGQLLVYCVNYFI	185	
Qy	171	ARLP----	LHL-GWRVMLAIGAVPSGLLALLVFCMPSPRWLVKGRLLADARAVLEK----	222
Dd	186	ARSGDASWLNTDGRYMFASECIPALLFLMLLYTPESPRLWMSRGKQEQAEGLRKIMG	245	
Qy	223	-TSATPEEAAERLADIKAAAGIPKGLDGDVVTVPGKEGGGELQVWKKLILSPTPAVRRI	281	
Dd	246	NLTAT-----	QAVQEIHKSLD-----HGRKTGG-----	RL 270
Qy	282	LLSAVG-----	LHFFQASGSDSVWQYSARLFKSAGITDDNKLGLGVTCAVGVTKTFFI	334
Dd	271	LMFGVGIVIGVMLSIFQQFGINVVLYYAPEVEFKTLGASTDIAL-L-QTIIIVGINLTFT	329	
Qy	335	LVATFLDRAGRRPLLLISTGGMIVSLICLGSLTVAGHPDTKVAVAVALCIASLTLSYI	394	
Dd	330	VLAIMTVDKFGRKPLQIIGALGMAIGMFSLGTAFY-----	TQAPGIVAL--LSMLFYV	380
Qy	395	AFSICLGPITGVYTSEIFPLQVRALGFVGVASNRVTSAVISMTFLSLSKAITI----	449	
Dd	381	ARFAMSWGVPVCWLLSEIFFPNAIRGKALAIATAVAAQWLANYFVSWTFPMDKNSWLVAHFH	440	
Qy	450	-GGSFFLYSGIAAVAVVFFFTCLPETRGRTLEEMGKLFGMPDPTGMAEE	496	
Dd	441	NGFSYWIYCGMGVLAALFMWKFVPETKGTLEELEALW-EPETKKTQQ	487	

RESULT 13

ARAE_BACSU	STANDARD;	PRT;	464 AA.
ID	ARAE_BACSU		
AC	P96710;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Arabinose-proton symporter (Arabinose transporter).		
GN	ARAE OR BSU33960.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1423;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=168;		
RX	MEDLINE=98044033; PubMed=9384377;		
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,		
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,		
RA	Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,		
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,		
RA	Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,		
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,		
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,		

CC	the promoter. L-arabinose acts as an inducer by inhibiting the
CC	binding of araR to the DNA, thus allowing expression of the gene.
CC	-!- SIMILARITY: Belongs to the sugar transporter family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; Z99121; CAB15401.1; -.
DR	EMBL; X98354; CAA66998.1; -.
DR	EMBL; Y12105; CAA72812.1; -.
DR	PIR; F69587; F69587.
DR	Subtilist; BG11907; araE.
DR	InterPro; IPR007114; MFS.
DR	InterPro; IPR005828; Sub_transporter.
DR	InterPro; IPR005829; Sug_transporter.
DR	InterPro; IPR003663; Sugar_transp.
DR	Pfam; PF00083; sugar tr; 1.

Db	298	QQFTGINAIIFYVPVLFSSLSGSSASSAALLN-TVVVGAVNVGSTMIAVLLSDKFGRRFLLI	356
Qy	352	ISTGGMIVSLICLGSLTVAGH-----HPDTKVAWAVALCIASTLSYIAFFSIGL	401
Db	357	--EGGITCCLAMLAAGITLGVVEFGQYGTEDLPHPVSAAGVLAV-ICI-----PIAGFAWSW	408
Qy	402	GPITGVVYTS EIFPLQVRALGFAGVAVSNRVTS AVISMTEFLSLSKAITIGGSPFLYSGIAA	461
Db	409	GPMGWLIPSEIFTLETRPAGTAVAVMGNFLESFVIGQAFVSMLCAMKF-GVFLFFAGWL	467
Qy	462	VAWVEFFTCLPETRGRRTLEMGKLF-----MPDTGMAEEAEEDAAAKEKVE	508
Db	468	IMVICAIFLLPETKGVPIERVQALYARHWFVKVMGPAAQEIIEAEDEKRVASQAIMK	525

Search completed: June 30, 2004, 18:36:01
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 18:36:04 ; Search time 55 Seconds
(without alignments)
2942.924 Million cell updates/sec

Title: US-10-051-902A-20
Perfect score: 2559
Sequence: 1 MASDELAKEVPRKKGNVY.....AEEAEDAAAKEKVVELPSSK 513

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1960.5	76.6	577	10 Q945E5	Q945e5 oryza sativ
2	1648.5	64.4	511	10 Q8LHC3	Q8lhc3 oryza sativ
3	1583	61.9	538	10 Q84QH3	Q84qh3 prunus cera
4	1515.5	59.2	509	10 Q84KI7	Q84ki7 prunus cera
5	1509	59.0	523	10 Q7XA50	Q7xa50 glycine max
6	1493	58.3	481	10 Q84RI1	Q84ri1 malus domes
7	1492.5	58.3	539	10 Q9LS92	Q9ls92 arabidopsis
8	1483.5	58.0	545	10 P93076	P93076 beta vulgar
9	1483.5	58.0	549	10 P93075	P93075 beta vulgar
10	1482	57.9	519	10 Q7XB36	Q7xb36 orobanche r
11	1479.5	57.8	511	10 Q9XIH7	Q9xih7 arabidopsis
12	1469.5	57.4	491	10 Q84RI2	Q84ri2 malus domes
13	1463.5	57.2	511	10 Q9XIH6	Q9xih6 arabidopsis
14	1420	55.5	513	10 Q9FQX3	Q9fqx3 apium grave
15	1346	52.6	493	10 O23213	O23213 arabidopsis
16	1316	51.4	508	10 Q9ZNS0	Q9zns0 arabidopsis

17	1274	49.8	506	10 Q8W2W8	Q8w2w8 oryza sativ
18	1274	49.8	506	10 Q7XFD6	Q7xfd6 oryza sativ
19	1274	49.8	574	10 Q9AUM9	Q9aum9 oryza sativ
20	1272.5	49.7	523	10 Q8RVQ2	Q8rvq2 apium grave
21	1176	46.0	547	10 Q9SKT9	Q9skt9 arabidopsis
22	1151	45.0	538	10 Q7XKF1	Q7xkf1 oryza sativ
23	1133.5	44.3	523	10 Q7X6M3	Q7x6m3 oryza sativ
24	1105	43.2	535	10 Q7XKF0	Q7xkf0 oryza sativ
25	668	26.1	457	16 P96742	P96742 bacillus su
26	635.5	24.8	245	10 Q8VZ80	Q8vz80 arabidopsis
27	629.5	24.6	580	10 Q9C757	Q9c757 arabidopsis
28	624.5	24.4	618	11 Q921A2	Q921a2 rattus norv
29	617	24.1	596	10 Q7XIZ0	Q7xiz0 oryza sativ
30	615.5	24.1	467	16 Q8CQA7	Q8cqa7 staphylococ
31	613	24.0	409	16 Q83EH4	Q83eh4 coxiella bu
32	610.5	23.9	581	10 Q9LKH1	Q9lkh1 mesembryant
33	610	23.8	544	10 Q93WT7	Q93wt7 olea europa
34	606	23.7	521	10 O22848	O22848 arabidopsis
35	602	23.5	546	10 Q93Z41	Q93z41 arabidopsis
36	600	23.4	509	10 Q8VZR6	Q8vzr6 arabidopsis
37	597	23.3	582	10 O23492	O23492 arabidopsis
38	595	23.3	560	10 Q9LFI3	Q9lfi3 arabidopsis
39	594.5	23.2	515	10 Q9LLD8	Q9lld8 arabidopsis
40	592.5	23.2	534	10 Q9LLE1	Q9lle1 nicotiana t
41	591.5	23.1	537	3 Q873D5	Q873d5 neurospora
42	589	23.0	555	10 Q9FXY8	Q9fxy8 mesembryant
43	588	23.0	517	10 Q9SX48	Q9sx48 arabidopsis
44	587	22.9	475	10 O48537	O48537 prunus arme
45	586	22.9	473	16 O34718	O34718 bacillus su

ALIGNMENTS

RESULT 1
Q945E5
ID Q945E5 PRELIMINARY; PRT; 577 AA.
AC Q945E5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sugar transporter.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu P., Chen Q., Huang G., Yi K.;
RT "Molecular cloning of putative sugar transporter in rice."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF416867; AAL14615.1; -.
DR Gramene; Q945E5; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS50850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 577 AA; 61178 MW; 95651621AE980CD4 CRC64;

RL Plant Physiol. 131:1566-1575 (2003).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Montmorency; TISSUE=Fruit;

RX Zhifang G., Loescher W.H.;

RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

RL EMBL; AY100638; AAM44082.1; -.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0008643; P:carbohydrate transport; IEA.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub_transporter.

DR InterPro; IPR003663; Sugar_transpt.

DR InterPro; IPR005829; Sug_transporter.

DR Pfam; PF00083; sugar_tr; 1.

DR PRINTS; PR00171; SUGTRNSPORT.

DR TIGRFAMS; TIGR00879; SP; 1.

DR PROSITE; PS00850; MFS; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

SQ SEQUENCE 538 AA; 58569 MW; ADDD03B983C6D358 CRC64;

Query Match 61.9%; Score 1583; DB 10; Length 538;

Best Local Similarity 65.1%; Pred. No. 9.3e-100;

Matches 310; Conservative 66; Mismatches 96; Indels 4; Gaps 2;

QY 12 PRKKGNVKYASICAILASMASVILGYDIGVMSGAAVYIKKDLNITDVQLEILIGILSLYS 71

Db 26 PKPKENKYAFACAILASMTSILLGYDIGVMSGAVIYIKKDLKVSDEIEVLVGIINLYS 85

QY 72 LFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGM 131

Db 86 LIGSAAAGRTSDWIGRRYTIIVLAGAIFAFGALLMGFAPNYAFMLFGRFVAGIGVGYALMI 145

QY 132 APVYTABISPAAARGFLTTTFPEVFINIGILLGYLSNFAPARLPLHLGWRVMLAIGAVPSG 191

Db 146 APVYTAEVSPASSRGFLTSFPPEVFINAGILFGYVSNYGFSKLPTHLGWRMLGVGAIPSI 205

QY 192 LLALLVFCMPESPRLWLKGLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGDV 251

Db 206 FLAIGVLAMPESPRLWVMOGRLGDKVLDKTSLSLEESKRLRGEIKEAAGIPEHCNDDI 265

QY 252 VTVPGKEQGGGELQVWKLLISPTPAVRRILLSAVGLHFFQQASGSDSVVQYSARLFKSA 311

Db 266 VEVKRSQG--QEVWKQLLRPTPAVRHILMCAVGLHFFQQASGIDAVVLSPRIFEKA 322

QY 312 GITDDNKLIGVTCVAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLICLGSGLTVA 371

Db 323 GITNPDHVLCTVAVGVFKTVFIFLVATFMLDRIGRRPLLLTSVAGMVFVTLACLGGLTII 382

QY 372 GHHPDTKWAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFVAVGNRV 431

Db 383 DHSGE-KIMWAIALSITWVLAIVAFSSIGMGPIITWVYSSEIFPLQLRAQGCISGVAVNRV 441

QY 432 TSAVISMTEFLSLSKAITIGGSFELYSGIAAVAVVFFFTCLPETRGRRTLEEMGKLFG 487

Db 442 VSGVLSMTFISLYKAITIGGAFFLFAAIAAVGWTFFFTMLPETQGRRTLEDMEVLFG 497

RESULT 4

Q84KI7 PRELIMINARY; PRT; 509 AA.

ID Q84KI7;

AC Q84KI7;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Sorbitol transporter.

OS Prunus cerasus.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.

OX NCBI_TaxID=140311;

RN SEQUENCE FROM N.A.

RP TISSUE=Fruit;

RX MEDLINE=22578918; PubMed=12692316;

RA Gao Z., Maourisset L., Lemoine R., Yoo S.-D., van Nocker S., Loescher W.;

RT "Cloning, Expression, and Characterization of Sorbitol Transporters from Developing Sour Cherry Fruit and Leaf Sink Tissues.";

RL Plant Physiol. 131:1566-1575 (2003).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Fruit;

RA Zhifang G., Loescher W.H., Lemoine R.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF482011; AAO39267.1; -.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0008643; P:carbohydrate transport; IEA.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub_transporter.

DR InterPro; IPR003663; Sugar_transpt.

DR InterPro; IPR005829; Sug_transporter.

DR Pfam; PF00083; sugar_tr; 1.

DR PRINTS; PR00171; SUGTRNSPORT.

DR TIGRFAMS; TIGR00879; SP; 1.

DR PROSITE; PS00850; MFS; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

SQ SEQUENCE 509 AA; 55407 MW; EA25ED065CE884C8 CRC64;

Query Match 59.2%; Score 1515.5; DB 10; Length 509;

Best Local Similarity 62.6%; Pred. No. 3.5e-95;

Matches 299; Conservative 72; Mismatches 102; Indels 5; Gaps 3;

QY 10 VEPRKGNVKYASICAILASMASVILGYDIGVMSGAAVYIKKDLNITDVQLEILIGILSL 69

Db 17 LKXKPKRNL-YAIGCAILASMTSILLGYDIGVMSGASIIYIKDLKISDVEVEILIGILNL 75

QY 70 YSLFGSFAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGVG 129

Db 76 YSLIGSAAAGRTSDWIGRRYTIIVFAGAIFFTGALLMGLATNYAFMLVGRFVAGIGVGYAL 135

QY 130 MIAPVYTABISPAAARGFLTTTFPEVFINIGILLGYLSNFAPARLPLHLGWRVMLAIGAVP 189

Db 136 MIAPVYNAEVSPASSRGALTSFPPEVFVFNIGILLGYVANYAFSGLPIDLGWRMLGVGVFP 195

QY 190 SGLLALLVFCMPESPRLWLKGLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDG 249

Db 196 SVILAVGVLSMPESPRLWVMOGRLGDKVLDKTSLSLEAQLRLADIKAAAGIPEHCVE 255

QY 250 DVTVPGKEQGGGELQVWKLLISPTPAVRRILLSAVGLHFFQQASGSDSVVQYSARLFK 309

Db 256 DVQVQPKHSHGE--EVWKELLHPTPPVRHILIAAIGFHFQQLSGIDALVLYSPRIFE 312

QY 310 SAGITDDNKLIGVTCVAVGVTKTFFILVATFLLDRAGRRPLLLISTGGMIVSLICLGSGLT 369

Db 313 KAGITDSSTLLLATVAVGFSKTIPTLVAIGFLDRVGRRRPLLLTSVAGMIASLLCLGTSLT 372

QY 370 VAGHHPDTKWAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFVAVGN 429

Db 373 IVDHETE-KMMWASVLCITWVLAIVGFFSIGMGPIAWVYSSEIFPLKLRAGQCCSMGTAVN 431

QY 430 RVTSAVISMTFSLSKAITIGGSFELYSGIAAVAVVFFFTCLPETRGRRTLEEMGKLFG 487

Db 432 RIMSGVLSMSPISLYKAITMGGTFELYAGIATVGWVFFFTMLPETQGRRTLEDMEVLFG 489

RESULT 5

Q7XA50 PRELIMINARY; PRT; 523 AA.

ID Q7XA50

AC Q7XA50;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)


```

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sorbitol-like transporter.
GN STP.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RA Dimou M., Aivalakis G., Katinakis P.;
RT "Carbon metabolism in developing roots and lateral roots of etiolated
RT Glycine max seedlings. Expression of GmSTP gene coding for a sorbitol-
RT like transporter.";
RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ563367; CAD91337.1; -.
DR EMBL; AJ563367; CAD91337.1; -.
SQ SEQUENCE 523 AA; 56489 MW; 3A586397262F1C4B CRC64;

Query Match 59.0%; Score 1509; DB 10; Length 523;
Best Local Similarity 61.0%; Pred. No. 9.9e-95;
Matches 303; Conservative 68; Mismatches 112; Indels 14; Gaps 3;

QY 1 MASDELAKAVE-----PRKKGNVKYASICAILMASVILGYDIGVMSGAAMYIK 50
Db 1 MTEGKLEAAEAHKTLDPPPKRKRNKYAFACAMLASMTSILLGYDIGVMSGAAMYIK 60

QY 51 KDLNITDVQLEILIGILSLYSLFGSFAGARTSDRIGRRLLTVFAAIVFFVGSLLMGFAVN 110
Db 61 RDLKVSDEQIEILLGIINLYSLIGSCLAGRTSDWIGRRYTTIGLGAIFLVGSTLMGFYPH 120

QY 111 YGMLMAGRFVAGVGVGGMIAPIVYTAETISPAASRGFLTTFPEVFINIGILLGYLSNFAF 170
Db 121 YSFLMCGRFVAGIGIGYALMIAPVYTAEVSPASSRGFLTSFPEVFINIGILLGYISNYGF 180

QY 171 ARLPLHLGWRVMLAIGAVPSGLLALLVFCMPESPRLWLVLKGLADARAVLEKTSATPEEA 230
Db 181 SKLTLLKVGWRMLGVGAIPSVVLTEGVLMPESPRLVLMRGLGEARKVNLKTSDSKEEA 240

QY 231 AERLADIKAAAGIPKGLDGDVVTPGKEQGGELQVWKKLILSPTPAVRRIILLSAVGLHF 290
Db 241 QLRLAEIKQAAGIPESCNDDVVQVKNQSGEG---VWKELFLYPTPAIRHIVIAALGIHF 297

QY 291 FQOASGSDSVVQYSARLFKSAGITDDNKLGLVTCAGVTKTFFILVATFLDLRAGRRPLL 350
Db 298 FQOASGVDVAVLYSPRIFEKAGITNDTHKLLATVAVGVKTVFILAATFLDRVGRRLPL 357

QY 351 LISTGGMIVSLICLGSGLTVAGHHPDTKVAVAVALCIASTLSYIAFESIGLGPITGVYTS 410
Db 358 LSSVGGMVLSTLTLAISLTVI-DHSERKLMWAVGSSIAMVLAYVATFSIGAGPITWVYSS 416

QY 411 EIFPLQVRALGFVAVASNRVTSAVISMTFLSLSKAITIGSGFFLYSGIAAVAVVFFFTC 470
Db 417 EIFPLRLRAQGAAGVAVNRVTSVAVSVMTFLSLTRAITIGGAFFLYCGIATVGWIFFYTV 476

QY 471 LPETRGRITLEEMGKLF 487
Db 477 LPETRGRITLEEMGSGF 493

RESULT 6
Q84RI1
ID Q84RI1 PRELIMINARY; PRT; 481 AA.
AC Q84RI1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sorbitol transporter.
GN SOT2.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

```

```

OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Mutsu; TISSUE=Fruit;
RA Zhifang G., Jayanty S., Beaudry R.M., Loescher W.H.;
RT "Cloning and characterization of apple fruit sorbitol transporter.";
RT Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY237401; AAC88965.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
SQ SEQUENCE 481 AA; 52566 MW; A3C20EE544A85C6E CRC64;

Query Match 58.3%; Score 1493; DB 10; Length 481;
Best Local Similarity 65.4%; Pred. No. 1.1e-93;
Matches 300; Conservative 61; Mismatches 92; Indels 6; Gaps 4;

QY 30 MASVILGYDIGVMSGAAMYIKKDLNITDVQLEILIGILSLYSLFGSFAGARTSDRIGRRLL 89
Db 1 MTSILLGYDIGVMSGASLFIKENLKISDVQVEINMGTLNLYSLIGSALAGRTSDWIGRRY 60

QY 90 TVVFAAIVFFVGSLLMGFAVNYGMLMAGRFVAGVGVGGMIAPIVYTAETISPAASRGFLT 149
Db 61 TIVLAGTIFFIGALLMGFAPNYAFLMFGRFVAGVGVGYALMIAPVYTAETISPAASRGFLT 120

QY 150 TPPEVFINIGILLGYLSNFAFARLPLHLGWRVMLAIGAVPSGLLALLVFCMPESPRLWLVL 209
Db 121 SPPEVFNIGILLGYVSNYAFSKLPIHLNWRIMLGVGAFPSVILAVGLAMPESPRLVLM 180

QY 210 KGLADARAVLEKTSATPEEAAERLADIKAAAGIPKGLDGDVVTPGKEQGGELQVWKK 269
Db 181 QGRGLDAKRVLQKTSIESIECQLRLDDIKEAAGIPKESNDDVVQVSKSHGEG--VWKE 237

QY 270 LILSPTPAVRRIILLSAVGLHFFQOASGSDSVVQYSARLFKSAGITD-DNKLGLVTCAGV 328
Db 238 LLLHPTPAVRHILIAALGIHFFEQSGSDSVVLYSPRIFEKAGITSYDHKLL-ATVAVGV 296

QY 329 TKTFEILVATFLDLRAGRRPLLILISTGGMIVSLICLGSGLTVAGHHPDTKVAVAVALCIA 388
Db 297 VKTICILVATVFLDKFGRRLPLLTLSVAGMVFSLSCLGASLTIVDQO-HGKIMWAIVLCIT 355

QY 389 STLSTYIAFFSIGLGPITGVYTS EIFPLQVRALGFVAVASNRVTSAVISMTFLSLSKAIT 448
Db 356 MVLNVAFFSIGLGPITWVYSS EIFPLQVRALGFVAVASNRVTSAVISMTFLSLSKAIT 415

QY 449 IGSFFFLYSGIAAVAVVFFFTCLPETRGRITLEEMGKLF 487
Db 416 IGAFFLYAGIAAVGVVFFYMLYPETQGRITLEDMEVLF 454

RESULT 7
Q9LS92
ID Q9LS92 PRELIMINARY; PRT; 539 AA.
AC Q9LS92;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

```


RP	SEQUENCE FROM N.A.	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
RC	TISSUE=Phloem;	DE	Sugar transporter like protein.
RX	MEDLINE=21149881; PubMed=112511106;	GN	AT4G36670.
RA	Noiraud N., Maurousset L., Lemoine R.;	OS	Arabidopsis thaliana (Mouse-ear cress).
RT	"Identification of a mannitol transporter, agmat1, in celery phloem.";	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
RL	Plant Cell 13:695-705(2001).	OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).	OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.
CC	-!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.	OX	NCBI_TaxID=3702;
DR	EMBL; AF215837; AAG43998.1; -.	RN	[1]
DR	GO; GO:0016021; C:integral to membrane; IEA.	RP	SEQUENCE FROM N.A.
DR	GO; GO:0005351; F:sugar porter activity; IEA.	RA	Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Schueller C.,
DR	GO; GO:0005215; F:transporter activity; IEA.	RA	Chalwatzis N.;
DR	GO; GO:0008643; P:carbohydrate transport; IEA.	RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR	InterPro; IPR007114; MFS.	RN	[2]
DR	InterPro; IPR005828; Sub_transporter.	RP	SEQUENCE FROM N.A.
DR	InterPro; IPR003663; Sugar_transpt.	RA	EU Arabidopsis sequencing project;
DR	InterPro; IPR005829; Sug_transporter.	RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR	Pfam; PF00083; sugar_tr; 1.	CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR	PRINTS; PR00171; SUGTRNSPORT.	CC	-!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR	TIGRFAMS; TIGR00879; SP; 1.	DR	EMBL; Z99708; CAB16808.1; -.
DR	PROSITE; PS50850; MFS; 1.	DR	EMBL; AL161589; CAB80333.1; -.
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.	DR	PIR; A85433; A85433.
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.	DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	Sugar transport; Transmembrane; Transport.	DR	GO; GO:0005351; F:sugar porter activity; IEA.
KW	Sugar transport; Transmembrane; Transport.	DR	GO; GO:0005215; F:transporter activity; IEA.
SQ	SEQUENCE 513 AA; 56013 MW; E08D655B9E08781 CRC64;	DR	GO; GO:0008643; P:carbohydrate transport; IEA.
	Query Match 55.5%; Score 1420; DB 10; Length 513;	DR	InterPro; IPR007114; MFS.
	Best Local Similarity 55.6%; Pred. No. 1.1e-88;	DR	InterPro; IPR005828; Sub_transporter.
	Matches 282; Conservative 91; Mismatches 116; Indels 18; Gaps 6;	DR	InterPro; IPR003663; Sugar_transpt.
		DR	InterPro; IPR005829; Sug_transporter.
QY	11 EPRKGNVKYASICAILASMASVILGYDIGVMSGAMVIKDLNITDVQLEILIGILSLY 70	DR	Pfam; PF00083; sugar_tr; 1.
Db	15 KFKPKRN-KYAFACALLASMNSILLGYDTGLSGASIYIKEDLHFSVDQIEIIGIINIIY 73	DR	PRINTS; PR00171; SUGTRNSPORT.
QY	71 SLFGSPAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVGYGGM 130	DR	TIGRFAMS; TIGR00879; SP; 1.
Db	74 SLGSAIAGRTSDWIGRRYTMVLAGIIFFLGAIFMGLATNFAFLMFGRFVAGIGVGYAMM 133	DR	PROSITE; PS50850; MFS; 1.
QY	131 IAPVYTAEISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWRVMLAIGAVPS 190	DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
Db	134 IAPVYTAEVAPSSSRGFLTSFPEVFINSVLLGYVSNFAFAKCPWLWLRMLGIGAFPS 193	DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
QY	191 GLLALLVFCMPESPRWLVLKGRLLADARAVLEKTSATPEEAERLADIKAAAGIPKGLDGD 250	DR	Sugar transport; Transmembrane; Transport.
Db	194 VALAIIVLYMPESPRWLVMQGRLEARTVLEKTSKBEAHQRLSDIKEAAGIDKDCNDD 253	KW	Sugar transport; Transmembrane; Transport.
QY	251 VVTVPKGEGGELQVWKLIILSPTPAVRRILLSAVGLHFFQOASGSDSVVQYSARLFKS 310	SQ	SEQUENCE 513 AA; 56013 MW; E08D655B9E08781 CRC64;
Db	254 VVQVPKRTK---DEAVWKELIHLPTKPVRHAAITGIGIHFFQOACGIDAVVVLSPRIFEK 310		Query Match 52.6%; Score 1346; DB 10; Length 493;
QY	311 AGITDDNKLGLGVTCAVGVTKTFILVATFLLDRAGRRPLLLISTGGMIVSLICLGSGLTV 370		Best Local Similarity 55.5%; Pred. No. 1.2e-83;
Db	311 AGIKNSKKLLATIAVGVCKTVFILLISTFQDKIGRRPLMLTSMGGMVIALFVLAGSLTV 370		Matches 268; Conservative 94; Mismatches 115; Indels 6; Gaps 5;
QY	371 --AGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFAVGAS 428	QY	6 LAKAVEPRKKNV-KYASICAILASMASVILGYDIGVMSGAMVIKDLNITDVQLEILI 64
Db	371 INKSHRTG---HWAGGLAIFTVYAFVSIFSSGMGPFAWVYSSEVFPRLRAQGSIGVAV 427	Db	1 MADQISGEKPGAGVNRFAHQCAIVASIVSIIFYGDTGVMSGAMVFIEDLKTNDVQIEVLT 60
QY	429 NRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAAVAVFFFTCLPETRGRTLERMKCLFG- 487	QY	65 GILSLYSLFGSPAGARTSDRIGRRLTVVFAAVIFFVGSLLMGFAVNYGMLMAGRFVAGVG 124
Db	428 NRGMSGIGMTFISMYKAMTIGGAPLLFAVVASIGVFMVMTFPTQGRNLEIIEILLFGS 487	Db	61 GILNLCALVGSLLAGRTSDIIGRYTIVLASILFMLGSLMGWPNYPVLLSGRCTAGLG 120
QY	488 -----MPDTGMAEEAEADAKEKV 507	QY	125 VGYGGMIAPVYTAEISPAASRGFLTTFPEVFINIGILLGYSNFAFARLPLHLGWRVMLA 184
Db	488 YFGWRKTLKDL-KAKEAAEAKSRESEV 513	Db	121 VGFALMVAPVYSABIATASHRGLLASLPHLCISIGILLGYIVNYFFSKLPMHIGWRMLG 180
		QY	185 IGAVPSGLLALLVFCMPESPRWLVLKGRLLADARAVLEKTSATPEEAERLADIKAAAGI- 243
		Db	181 IAAVPSLVLAFGILKMPESPRWLIMQGRLEKEKEILELVNSPPEEAELRFQDIKAAAGID 240
QY		QY	244 PKGLDGDVVTVPGKEQGGELQVWKLIILSPTPAVRRILLSAVGLHFFQOASGSDSVVQY 303
Db		Db	241 PKCVD-DVVKMEGKKTHGE--GVWKELILRPTPAVRRVLLTALGIIHFFQHASGIEAVLLY 297
QY		QY	304 SARLFKSAGITDDNKLGLGVTCAVGVTKTFILVATFLLDRAGRRPLLLISTGGMIVSLIC 363
Db		Db	298 GPRIFKKAGITTKDKLFLVTIGVGIMKTTFTATLLEDKVGRRKLLLTSGVMVIALTM 357
		QY	364 LGSGLTVAGHHPDTKVAWAVALCIASTLSYIAFFSIGLGPITGVYTSEIFPLQVRALGFA 423
RESULT 15		Db	358 LGFGLTMA-QNAGGKLAWALVLSIVAAYSFVAFFSIGLGPITVWVYSSEVFPKLKRAQGAS 416
O23213	PRELIMINARY; PRT; 493 AA.	QY	424 VGVASNRVTSAVISMTFLSLSKAITIGGSFFLYSGIAAAVAVFFFTCLPETRGRTLLEEMG 483
AC	O23213;	Db	417 LGVAVNRVNMATVSMFSLSLTSAITTGCAFFMFAGVAAVAMNPPFFLLPETKGSLEEIE 476
DT	01-JAN-1998 (TrEMBLrel. 05, Created)		
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)		

Qy 484 KLF 486
Db 477 ALF 479

Search completed: June 30, 2004, 18:42:52
Job time : 69 secs